INVESTIGATION OF THE REMOVAL OF CONTAMINANTS OF EMERGING CONCERN AND THE MICROBIOME IN BIOFILTERS

A DISSERTATION

SUMMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

BEN MA

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

RAYMOND M. HOZALSKI, ADVISOR

APIRL 2019

ACKNOWLEDGMENTS

First, I would like to thank my advisor, Dr. Raymond Hozalski, for being a great mentor, role model and friend. He has provided me with constant supports and guidance on conducting research, interpreting results, writing papers and reports, presenting in public, and pursuing an academic career in the future. His commitment to my success and well-being helped me through my study in the graduate school. It has been my privilege working with him in the past four and half years.

Next, I would like to thank Dr. William Arnold and Dr. Timothy LaPara for helping the method development and interpretation of results in my research projects, and serving in my dissertation committee. I would like to thank other committee members, Dr. Matt Simcik and Dr. Sebastian Behrens for reviewing this thesis and providing valuable feedbacks. Finally, I would like to thank my colleagues for providing suggestions in my research.

Financial support for this work was provided by the City of Minneapolis and Water Research Foundation (WRF 4669). Thanks to Minneapolis Water Treatment and Distribution Services for their support during the operation of the pilot-scale biofilters.

Many thanks to my friends and my parents for helping me through some difficult times. Finally, I would like to thank everything happened to me in the past four and half years, good and bad, made me who I am.

ABSTRACT

The removal and fate of contaminants of emerging concern (CECs) in water treatment systems is of interest given the widespread occurrence of CECs in water supplies. Biofiltration, which granular media filters are inhabited by viable bacteria, has the potential of providing long-term, sustainable CECs removal in drinking water treatment. Bacterial communities in biofilters can be beneficial through biodegradation of contaminants but also pose potential risks by harboring and releasing detrimental microbes into water distribution systems. In this work, the removal of eight CECs, including atenolol, atrazine, carbamazepine, fluoxetine, gemfibrozil, metolachlor, sulfamethoxazole and tris(2-chloroethyl) phosphate, was investigated in pilot-scale granular activated carbon (GAC)-sand and anthracite-sand biofilters. The effects of water quality and engineering decisions on the biofilter microbiome and the effect of biofilters on the microbiome in filter effluent were evaluated in the aforementioned pilot-scale biofilters. In addition, the geographic patterns of the biofilter microbiome were investigated by sampling filter media from full-scale biofilters at fourteen treatment plants throughout North America. The CECs concentrations in the filter influent and effluent were determined using liquid chromatograph tandem mass spectrometry, and the bacterial abundance and community composition in the biofilters were determined using real-time quantitative polymerase chain reaction (qPCR) and Illumina HiSeq high-throughput sequencing of PCR amplicons. GAC-sand biofilters provided superior CECs removal for all compounds (mean removal efficiencies: 49.1-94.4%) compared to anthracite-sand biofilters (mean removal

efficiencies: 0-66.1%) due to a combination of adsorption and biodegradation. Adsorption was determined to be the dominant removal mechanism for most selected CECs in GAC-sand biofilters. A multiple linear regression based empirical relationship considering water quality, engineering decisions, and CECs chemical properties was developed to predict CECs removal in the GAC-sand biofilters. The microbiome in the pilot-scale and full-scale biofilters contained genera that are commonly found in the freshwater environments and water distribution systems, such as *Limnohabitans*, *Flavobacterium*, *Nitrospira*, and, *Hydrogenophaga*. The microbiome in the pilot-scale biofilters exhibited temporal variations, and varied with media type (GAC vs. anthracite), backwash strategy (chloraminated vs. non-chloraminated), and bed depth. The pilot-scale biofilters effectively removed biomass (~70%) from the water, but only marginally impacted the microbiome in the filter effluent. Significant inter-filter variations were observed in the full-scale biofilter investigation that followed a weak but highly significant distance-decay relationship. The water quality characteristics exhibited a stronger influence on the microbiomes in the full-scale biofilters than the geographic distance according to a multiple regression on matrix analysis. *Nitrosomonas oligotropha*-like ammonia oxidizing bacteria (AOB) were generally more abundant than ammonia oxidizing archaea in the full-scale and pilot-scale biofilters. The ratios of nitrite oxidizing bacteria to AOB exceeded the theoretical ratio for conventional two step nitrification in most full-scale biofilters (12 of 14 biofilters) and in the pilot-scale biofilters for most of the operation. This work should be beneficial to drinking water treatment facilities in improving CECs removal performance using

biofilters, as well as to environmental engineers and scientists in understanding the microbiome in biofilters.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Anthropogenic activities have resulted in the intentional and unintentional release of contaminants to waters that serve as drinking water sources. Contaminants of emerging concern (CECs) are defined as a group of synthetic or naturally occurring chemicals that have been only recently detected in the environment using advanced analytical methods, such as liquid chromatography tandem mass spectrometry. CECs include a wide variety of chemical classes, such as insecticides and herbicides, pharmaceuticals and personal care products, and industrial and commercial chemicals, and have very different chemical structures, polarities, biodegradabilities and molecular weights. Concentrations of CECs in surface water supplies commonly range from 1 to 1000 ng/L [1]. Some CECs are persistent in drinking water treatment systems and have been detected in the finished water [2], which may have adverse effects on human health and negative consequences in terms of public perception.

There are some options to remove CECs in water treatment. Powder and granular activated carbon (PAC and GAC) have been used for sorption of CECs [3, 4]. Advanced oxidation processes (AOPs, including ozonation and UV) and membrane filtration (e.g., nanofiltration and reverse osmosis) are capable of degradation and rejection various CECs in water treatment processes [5–7]. Despite the effective CECs removal, these methods are

normally cost/energy intensive and require additional procedures to maintain the removal efficiencies (e.g., membrane cleaning for fouling control). Biofiltration, in which granular media filters are inhabited by viable microorganisms, has the potential of providing long-term, sustainable CECs removal in drinking water treatment with reasonable cost and energy consumption. This process has been used to remove natural organic matter [8, 9], taste-and-odor compounds [10] and other trace organic contaminants [11] via biodegradation. Biofiltration is not a new concept, because biofilters operated at low filtration rates, termed slow sand filters, have been used for water treatment since the 1800s [12]. Certainly, biodegradation alone may not be able to treat all dissolved contaminants, because some contaminants may be recalcitrant or biodegraded at a rate too slow to be removed in the relatively short hydraulic residence time of a rapid biofilter. Thus, biological treatment in a GAC filter is advantageous in that it provides the dual benefit of sorption and biodegradation.

The bacterial community on biofilter media can be beneficial in removing various contaminants through biodegradation. As a part of the source-to-tap microbiome, however, this community may also pose potential risks by harboring and releasing detrimental microbes into water distribution systems. There were a few studies concerning the bacterial community in biofilters. Bacteria like *Alphaproteobacteria*, *Gamaproteobacteria*, *Nitrospira* (class level) were consistently detected in full-scale biofilters [13, 14]. The bacterial community on filter media exhibited temporal variations and may help shaping the microbiome in water distribution systems [13–15]. Despite previous work, detailed

understanding of the variations and dynamics in biofilter microbiome is still in need as it may help improve biofilter performance and manage potential risks. Furthermore, it answers the ecological question of nurture versus nature. Namely, which factors have a stronger influence on the bacterial communities that develop in engineered biofilters, 'nature' factors such as the source water microbiome and water quality, or engineering decisions (i.e., 'nurture' factors) such as the choice of filter media, backwash conditions, and prefiltration treatment operations.

The three main objectives of this work were as follows: (1) determine the ability of biofilters to remove eight selected CECs; (2) evaluate the dynamics and effects of water quality and operating conditions on the microbiome in biofilters, and the impact of biofilters on the microbiome in filter effluent; and (3) investigate biogeographic patterns of the microbiome in full-scale biofilters across North America. The research entailed operation of eight pilot-scale biofilters in parallel for over 18 months at a full-scale water treatment plant. Six biofilters (three GAC-sand and three anthracite-sand) were fed a cocktail of eight CECs and two additional GAC-sand biofilters were operated without CECs dosing. In addition, media from full-scale biofilters at fourteen water treatment plants throughout North America was sampled to characterize their microbiomes.

In Chapter 2, the removal of eight CECs, including atenolol, atrazine, carbamazepine, fluoxetine, gemfibrozil, metolachlor, sulfamethoxazole, and tris(2-chloroethyl) phosphate, was investigated in the pilot-scale GAC-sand and anthracite-sand biofilters. The relative

contribution of sorption and biodegradation in the GAC-sand biofilters was determined by using the anthracite-sand biofilters as non-sorptive controls (i.e., biodegradation only). The effects of hydraulic loading rate or empty bed contact time (EBCT), influent CECs concentration, and throughput were evaluated. Finally, a regression analysis was performed to develop an empirical relationship between CECs removal in GAC-sand biofilters and parameters that were likely to affect CECs removal, including biodegradation rates, EBCT, influent CECs concentration, charge of CECs, influent dissolved organic carbon (DOC) concentration, and water temperature. The results of this chapter have been published [11]. It is reproduced in part with permission from the publisher (2019 Elsevier). Dr. Raymond Hozalski and Dr. William Arnold contributed to the experimental design, method development, and interpretation of results.

In Chapter 3, the dynamics and variations in the biofilter microbiome were investigated by sampling the aforementioned pilot-scale biofilters. The effect of water quality (e.g., temperature and DOC), bed depth, media type (GAC-sand vs. anthracite-sand), CECs addition, and backwash strategy (chloraminated vs. non-chloraminated) on bacterial abundance and community composition in the biofilter were determined. The bacterial abundance and community composition in the pre- and post-filtration water samples were also evaluated to determine the influence of biofilters on the microbiome in filter effluent. This chapter represents a draft manuscript that was prepared for submission to Microbiome journal. Dr. Raymond Hozalski and Dr. Timothy LaPara contributed to the experimental design, method development, and interpretation of results.

In Chapter 4, the bacterial abundance and community composition of full-scale biofilters at fourteen drinking water treatment facilities across North America were investigated to determine the effects of geographic location (i.e., inter-facility distance), contemporaneous water quality characteristics (e.g., pH, temperature, and influent DOC concentration), and engineering decisions (e.g., media type and prefiltration ozonation application) on the biofilter microbiomes. A multiple regression on matrix (MRM) analysis was performance to identify the relative importance of each abovementioned factor. This chapter represents a draft manuscript that was prepared for submission to the ISME journal. Dr. Raymond Hozalski and Dr. Timothy LaPara contributed to the experimental design, method development and interpretation of results. Ashely Evans coordinated the sample collection and provided the water quality data.

In Chapter 5, the abundance of nitrifying microorganisms, including ammonia oxidizing bacteria (AOB) and archaea (AOA), and nitrite oxidizing bacteria (NOB), was investigated in the aforementioned full-scale and pilot-scale biofilters. Nitrifying microorganisms are of interest because they indicate nitrification potentials of the biofilters. Furthermore, the presence of nitrifying microorganisms in the biofilters may lead to chloramine decay in chloraminated water distribution systems if these microorganisms are shed from the biofilters into the filter effluent and survive the subsequent disinfection procedures [16, 17]. The abundance of AOA and AOB were quantified by using real-time quantitative polymerase chain reaction (qPCR) targeting the ammonia monooxygenase subunit A

(*amoA*) gene in archaea and bacteria, respectively. The composition of nitrifying bacterial communities was investigated by sequencing the V3 region of the 16S rRNA gene on the Illumina HiSeq platform. The potential biases of qPCR assays were evaluated by comparing the qPCR results and DNA sequencing results. Finally, the ratio of NOB to AOB were investigated. This chapter represents a draft manuscript that was prepared for submission to Applied and Environmental Microbiology journal. Dr. Raymond Hozalski and Dr. Timothy LaPara contributed to the experimental design, method development and interpretation of results.

Taken as a whole, this thesis provides comprehensive investigation of the CECs removal and the microbiome in biofilters, which should be beneficial to drinking water treatment facilities in improving CECs removal performance using biofilters, as well as to environmental engineers and scientists in understanding the microbiome in biofilters. The conclusions of this work are provided in Chapter 6.

Chapter 2

Investigation of the removal of contaminants of emerging concern in biofilters.

The removal and fate of contaminants of emerging concern (CECs) in water treatment systems is of interest given the widespread occurrence of CECs in water supplies and increase in direct potable reuse of wastewater. In this chapter, CECs removal was investigated in pilot-scale biologically-active granular activated carbon (GAC)-sand and anthracite-sand filters under different hydraulic loading rates and influent CECs concentrations over a 15-month period. Eight of the most commonly detected compounds in a survey of CECs occurrence in drinking water were selected for this study: atenolol, atrazine, carbamazepine, fluoxetine, gemfibrozil, metolachlor, sulfamethoxazole and tris(2-chloroethyl) phosphate (TCEP). GAC-sand biofilters provided superior CECs removal for all compounds (mean removal efficiencies: 49.1-94.4%) compared to anthracite-sand biofilters (mean removal efficiencies: 0-66.1%) due to a combination of adsorption and biodegradation. Adsorption was determined to be the dominant removal mechanism for most selected CECs, except fluoxetine, which had the greatest biodegradation rate constant (0.93 ± 0.15 min^{-1} at 20-28 ^{o}C). The mean removal efficiency decreased by 16.5 % when the loading rate increased from 2 to 4 gpm/ft^2 (4.88 to 9.76 m/h). A significant reduction in CECs removal was observed after 100,000 bed volumes when the influent CECs concentration was low (100-200 ng/L), whereas no significant reduction

was observed during spike dosing (1000-3000 ng/L). A regression analysis suggested that biodegradation rate, hydraulic loading rate, influent CECs concentration, throughput, influent dissolved organic carbon (DOC) concentration, and CEC charge are important parameters for predicting CEC removal performance in GAC-sand biofilters.

2.1 Introduction

Contaminants of emerging concern (CECs) include a wide variety of chemical classes, such as insecticides and herbicides, pharmaceuticals and personal care products, and industrial and commercial chemicals. Based on a survey by Benotti et al. (2009), at least one CEC was detected in the source water of nineteen different U.S. drinking water utilities serving more than 28 million people. The CECs are present in source waters due to agricultural activity, industrial effluents, urban runoff, and municipal wastewater treatment plant effluent. It is not desirable to have these biologically-active compounds present in finished drinking water because there are possible adverse effects on human health and negative consequences in terms of public perception.

Biofiltration, a process in which granular media filters are inhabited by viable bacteria, has been used for water treatment since the 1800s [12]. Innovative engineering is not required to establish a biofilter, because most granular media filters are presumed to be biologically active regardless of upstream disinfection. Previous studies demonstrated that biofilters not only effectively remove particles but also remove a wide range of dissolved contaminants including ozonation byproducts [9, 18], halogenated disinfection byproducts (e.g., haloacetic acids; Zhang et al., 2009), geosmin [10, 20], and methylisoborneol [20].

Certainly, biodegradation alone may not be able to treat all dissolved contaminants, because some compounds may be recalcitrant or biodegrade at a rate too slow to affect the

CEC concentration in the short hydraulic residence time of a rapid biofilter. Thus, biological treatment in a granular activated carbon (GAC) filter is advantageous in that it provides the dual benefit of sorption and biodegradation. Furthermore, biological activity has been shown to substantially increase GAC bed life for some compounds, such as geosmin [10].

There are a few reports in the literature concerning the removal of CECs in drinking water treatment biofilters [5, 21–24]. Conventional sand or anthracite-sand biofilters can achieve 50-80% removal for biodegradable compounds, but little or no removal for compounds that typically are persistent in the environment (e.g., atrazine and carbamazepine) [21, 23, 24]. GAC biofilters, however, provided more than 90% removal of atrazine and carbamazepine over the first 200 days of operation [22]. Despite this prior work, there are several important knowledge gaps including: (1) the relative roles of sorption and biodegradation in CEC removal by GAC-sand biofilters and (2) the performance of biofilters when operated for both filtration and CEC removal over a wide range of conditions, including different empty bed contact times (EBCTs) and/or influent CEC concentrations. In this study, six pilot-scale biofilters (three with GAC-sand and three with anthracite-sand) were operated in parallel for 15 months at a full-scale water treatment plant to (1) evaluate the CEC removal performance of GAC-sand biofilters; (2) determine the relative contributions of sorption and biodegradation; and (3) assess the influence of operational conditions, including loading rate/EBCT, influent CECs concentration and throughput.

2.2 Methods

2.2.1 Pilot-scale biofiltration system design and operation

A pilot-scale biofiltration system (Figure 2-1) was constructed at the Minneapolis Water Treatment and Distribution services (MWTDS) Fridley facility, in Fridley, Minnesota, USA. MWTDS provides 57 million gallons per day ($215{,}000\ m^3/day$) of water to nearly a half-million customers in the city of Minneapolis and neighboring suburban communities. The Fridley facility treats the water by lime softening with coagulant addition, flocculation, primary sedimentation, recarbonation, secondary sedimentation, and anthracite-sand filtration. Water from the full-scale recarbonation chambers was used as the feed water for the pilot plant (Fig. 2-1). The pH of the water was adjusted to ~8.5 using hydrochloric acid and then ferric chloride was added as the primary coagulant at a dose of 2 mg/L as Fe. The water was settled for 1 to 2 hours (depending on the loading rate) and then spiked with CECs before entering the biofilter columns. Photos and schematic diagrams of the pilot-scale biofiltration system are provided in Appendix A

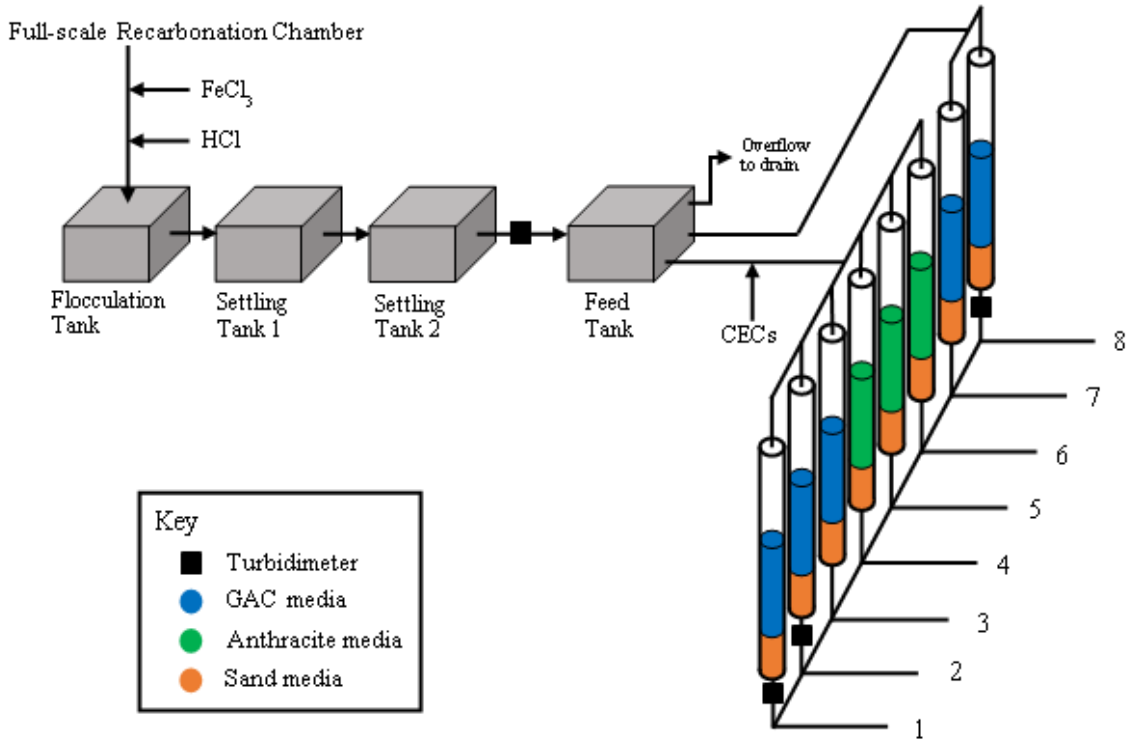

Figure 2-1. Schematic diagram of the pilot-scale biofiltration system at MWTDS. (The diagram is not drawn to scale).

Although eight biofilter columns were operated in parallel, only six biofilter columns are relevant for this chapter because the feed water to other two was not dosed with CECs. Each column was constructed from polycarbonate tubes (ID= 3.75 inches (9.53 cm), total height=144 inches (3.66 m)). Three columns were packed with 20 inches (0.51 m) of GAC (Calgon F300) over 10 inches (0.25 m) of sand and three columns were packed with 20 inches (0.51 m) of anthracite over 10 inches (0.25 m) of sand. The three anthracite-sand biofilter columns served as non-sorptive controls.

The pilot-scale biofiltration system was operated for 19 months from July 2015 to January 2017. The water temperature was not controlled and ranged from 3 to 28 ^{o}C. The filters were operated at hydraulic loading rates of 2 or 4 gpm/ft^2 (4.88 or 9.76 m/h), which provided 8.4 min or 4.2 min of empty bed contact time (EBCT), respectively. Backwash was performed when the headloss was above 90% of the maximum value permitted by the system (60 inches (1.53 m)). The columns were backwashed using 1.3 scfm (36.8 L/min) of air for 5 minutes followed by 10 minutes of water wash at 15 gpm/ft^2 (36.66 m/h) to achieve a 30% bed expansion.

Environmentally relevant amounts of CECs (100-200 ng/L) were continuously dosed into the filter influent using a syringe pump (NE-1600, New Era Pump Systems, Inc.; Farmingdale, NY) starting October 2015. The biofilters also were challenged periodically with sudden increases in CECs concentration (1000-3000 ng/L) for 4-6 hours to investigate the potential for breakthrough under extreme contaminant loading conditions. The details of CEC dosing are provided in the Table B-1. A 2×2×2 factorial design was used, with two loading rates (2 or 4 gpm/ft^2 (4.88 or 9.76 m/h)), two temperature ranges (low: 3 to 10 ^{o}C and high: 20 to 28 ^{o}C), and two CEC concentrations (low: 100 to 200 ng/L and high: 1000 to 3000 ng/L). Six of the eight different sets of conditions were repeated for a total of fourteen sampling periods throughout the experiment, with two sampling events per period. The same columns were used continuously throughout the study and exposed to the different conditions sequentially as described in Table 2-1. A total of 616 samples were collected for CECs analysis over the entire course of the study.

Filter influent samples were collected daily during weekdays and analyzed for temperature and pH. Pre- and post-filtration water samples were collected weekly for analysis of dissolved organic carbon (DOC) concentration. Biofilter media samples were collected monthly from the top layer of the biofilters for analysis of microbial activity.

Table 2-1. Sampling schedule and operational conditions of the pilot-scale biofiltration system. Temperature: Low (3-10 ^{o}C); High (20-28 ^{o}C); Influent CECs concentration: Low (100-200 ng/L); High (1000-3000 ng/L).

Sampling Period		Temperature	Loading Rate (gpm/ft^2)	Influent CECs concentration*	Throughput (as bed volume)
I	12/01/15-3/1/16	Low	2	Low	10286-21086
II	3/2/16 & 3/8/16	Low	2	High	22886-24429
III	3/9/16-3/31/16	Low	4	Low	33171-35229
IV	4/1/16 & 4/5/16	Low	4	High	36771-38829
V	6/29/16-7/14/16	High	2	Low	60686-62486
VI	7/15/16 & 7/19/16	High	2	High	64800-65829
VII	7/20/16-8/1/16	High	4	Low	68914-70971
VIII	8/2/16 & 8/5/16	High	4	High	73029-74571
IX	8/26/16-9/6/16	High	2	Low	79971-81514
X	9/7/16 & 9/9/16	High	2	High	83057-83571
XI	9/10/16-9/29/16	High	4	Low	89229-92829
XII	9/30/16 & 10/5/16	High	4	High	94371-96942
XIII	12/15/16-12/27/16	Low	2	Low	115200-116743
XIV	12/28/16 & 1/3/16	Low	2	High	118543-120086

*See Table B-1 for specific dosing concentrations.

2.2.2 CECs

Eight CECs were selected for this study: atenolol, atrazine, carbamazepine, fluoxetine, gemfibrozil, metolachlor, sulfamethoxazole, and tris (2-chloroethly) phosphate (TCEP). Details concerning the supplier and purity of each compound are provided in the Table B-2. These CECs were selected because they have been detected frequently in finished drinking water [1] and because they represent a wide range of uses and chemical characteristics. Aerobic biodegradation probabilities of the selected CECs were computed based on chemical structure using the MITI model [25] in BIOWIN6. The properties of the selected CECs are summarized in Table 2-2.

Table 2- 2. Chemical and biological properties of the selected CECs.

Compound	Use	Molecular Formula	Water solubility (mg/L)[b]	log D_{ow} (pH 8.5)[a]	BIOWIN6[c]	Charge behavior (pH 8.5)
Atenolol	Beta blocker	$C_{14}H_{22}N_2O_3$	13300	-1.04	0.2349	Positive
Atrazine	Herbicide	$C_8H_{14}ClN_5$	34.7	2.61	0.0000	Neutral
Carbamazepine	Mood stabilizer	$C_{15}H_{12}N_2O$	112	2.45	0.0364	Neutral
Fluoxetine	Antidepressant	$C_{17}H_{18}F_3NO$	17.7	2.73	0.0000	Positive
Gemfibrozil	Lipid regulator	$C_{15}H_{22}O_3$	10000	0.77	0.6123	Negative
Metolachlor	Herbicide	$C_{15}H_{22}ClNO_2$	530	3.13	0.0153	Neutral
Sulfamethoxazole	Antibiotic	$C_{10}H_{11}N_3O_3S$	610	-1.45	0.0060	Negative
TCEP	Flame retardant	$C_6H_{12}Cl_3O_4P$	7820	1.44	0.0193	Neutral

Note:

a. log D_{ow} (pH 8.5) (octanol-water partitioning coefficient at pH 8.5) is calculated from equation:

$$D_{ow} = \frac{K_{ow}}{1 + 10^{(pH - pKa)}}$$

in which K_{ow} values are estimated using EPI Suite™ (Temp: 20 °C). pKa values are estimated from ChemAxon.

b. Water solubility are estimated using EPI Suite™ WSKOW model.

c. Biodegradability are estimated using EPI Suite™ BIOWIN v4.10, MITI non-linear model. The values of BIOWIN6 represent the probabilities of readily aerobic biodegradation.

2.2.3 Batch sorption experiments

Batch sorption experiments were performed to determine the GAC sorption capacities for the CECs according to the procedure of Graham et al. (2000). Briefly, pre-combusted amber glass bottles were filled with 500 mL of carbonate-buffered (0.1 M) ultrapure water (pH 8.5) and then spiked with a cocktail of the target CECs (2 mg/L each). Activated carbon (AC) media was added into the bottles (0 to 312.5 mg AC/L) to initiate the adsorption experiment. The initial and equilibrium aqueous CEC concentrations were measured via high performance liquid chromatography (HPLC) with UV-Vis detection, except for TCEP, where liquid chromatography tandem mass spectrometry (LC-MS/MS) was used. Additional batch experiments were performed in a similar manner but with Mississippi River natural organic matter (NOM; International Humic Substances Society) added. The detailed experimental procedures and analytical methods are provided in the Appendix A.

The data from the batch sorption tests were fit to the Langmuir isotherm model using non-linear regression (IBM SPSS Statistics 22). When the aqueous concentration (C) is low, the Langmuir isotherm equation can be simplified to a linear isotherm as shown in Eq. 2-1 (by assuming $1 + K_l \times C \cong 1$):

$$q = \frac{K_l \times q_m \times C}{1 + K_l \times C} = K_d \times C \qquad (2-1)$$

where: q is the amount of CECs sorbed on GAC media (mg CECs/mg AC), K_l is the Langmuir constant (L/mg CECs), q_m is maximum sorption capacity (mg CECs/mg AC), C

is the aqueous CECs concentration, K_d is the carbon-water distribution coefficient (L/mg AC).

2.2.4 Analytical methods

The pH and temperature of the filter influent samples were measured using a HQ411D laboratory pH meter (Hach Company; Loveland, CO). The DOC concentration in the pre- and post-filtration water samples was determined according to EPA method 415.3 (American Public Health Association, 1998). To assess microbial activity levels on the filter media, the adenosine triphosphate (ATP) concentration of was determined immediately after sample collection using LuminUltra test kits according to the manufacturer's directions.

For measurement of CECs, quadruplicate filter influent samples and triplicate effluent samples (500 mL) from each biofilter column were collected in pre-combusted amber glass bottles and transported on ice to the laboratory. Upon arrival, all samples were filtered through glass fiber filters (0.7 μm) acidified to pH 2.0 with sulfuric acid, and then spiked with eight surrogate compounds (Table B-2) each at a concentration of 10 ng/L. To verify the relative recovery, one filter influent and one filter effluent sample from each column were spiked with 200 ng/L CECs as a matrix spike. Then, all samples were subjected to solid phase extraction (SPE) using 500 mg Oasis HLB cartridges following the protocol of Vanderford and Snyder (2006) and Trenholm et al. (2006) [27, 28]. After loading the SPE

cartridge, it was dried under vacuum (20 mmHg) for 150 seconds and eluted with 5 mL methanol followed by 5 mL 90% methyl tert-butyl ether / 10% methanol (v/v), and then the combined extracts were concentrated via nitrogen blowdown and solvent exchanged into 5 mM ammonium acetate solution (500 µL) for analysis via LC-MS/MS. LC-MS/MS method details are provided in the Appendix B.

Identification of potential microbial transformation products predicted using the University of Minnesota Biocatalysis / Biodegradation Database Pathway Prediction System [29] was attempted using ion trap mass spectrometry following the protocol of Helbling et al. (2010) [30].

2.2.5 Data analysis

Biodegradation rate constants were calculated by assuming the biofilter column is a plug-flow reactor and the biodegradation kinetics were pseudo-first-order according to Eq. 2-2:

$$k_{bio} = -\frac{1}{t} \times \ln(\frac{C_{eff}}{C_{inf}}) \qquad (2-2)$$

where: k_{bio} is the pseudo-first-order rate constant (min^{-1}), $t = \varepsilon \times EBCT$, is the hydraulic residence time (min) (ε=0.37, filter bed porosity), C_{eff} is the filter effluent concentration (ng/L), and C_{inf} is the filter influent concentration (ng/L).

The cumulative absorbent-phase concentration (q_c) for each CEC on the GAC media was calculated to assess the extent of bed capacity utilized over time. Because ATP concentrations on the filter media were similar for the GAC-sand and anthracite-sand biofilters (Fig. B-1), the contribution of biodegradation to CEC removal within the GAC-sand biofilters was assumed to be the same as in the anthracite-sand biofilters. Thus, q_c was calculated by subtracting CECs removal in the anthracite-sand biofilters from the mass removed in the GAC-sand biofilters according to the approach of Scharf et al. (2010):

$$q_c = \frac{\sum Q\left[\left(C_{inf} - C_{eff\_GAC}\right) - \left(C_{inf} - C_{eff\_Anth}\right)\right]t}{M_{GAC}} \qquad (2-3)$$

where: q_c is the cumulative CEC concentration on the GAC media (ng CEC/g GAC), Q is the flow rate (L/min), M_{GAC} is mass of GAC filter media (g), C_{inf} is the mean influent CEC concentration (ng/L), C_{eff_GAC} is the mean GAC-sand biofilter effluent CEC concentration (ng/L), C_{eff_Anth} is the mean anthracite-sand biofilter effluent CEC concentration, t is the biofilter operating time since CEC dosing began (min).

The maximum sorption capacity of the GAC media in the biofilters (q_e) was computed by inputting the biofilter influent CEC concentration into the sorption isotherm equation (Eq. 2-1) for each CEC ($q_e = K_d C_{inf}$). The value of q_c/q_e was computed to represent the fraction of occupied sorption sites on the GAC media in the biofilters.

The software program AdDesignS (Mertz et al., 1994), based on the pore and surface diffusion model (PSDM), was used to simulate the removal of CECs due to adsorption in

the GAC-sand biofilters over the course of the experiment. The isotherm parameters for the simulation were determined from the batch sorption experiments. Other parameters for the simulation were determined following AdDesignS manual recommendations (Mertz et al., 1994).

The statistical software IBM SPSS Statistics 22 was used to perform all statistical analyses. Paired and unpaired one-tailed two-sample t-tests were used to determine if the effects of hydraulic loading rate and throughput were statistically significant. A One-tailed one-sample t-test was used to determine if the mean of k_{bio} was significantly greater than zero. Pearson's correlation test was used to determine if a statistically significant correlation existed between k_{bio} and BIOWIN6.

2.2.6 Multiple linear regression analysis

A multiple linear regression analysis was performed to develop an empirical relationship between GAC-sand biofilter performance and parameters that were likely to affect CECs removal. A total of 220 CECs removal efficiency values were used in the analysis (14 sampling periods × 2 sampling events per period × 8 different CECs analyzed per sample, four removal efficiencies were excluded due to poor surrogate recovery). Operational and environmental factors that were likely to have a significant effect on CECs removal according to the literature [22, 31–33] were considered in the regression analysis: biodegradation, EBCT, influent CECs concentration, throughput, hydrophobicity and

charge of CECs, influent DOC concentration and water temperature. The coefficients (B_i) of the following first order empirical relationship were solved for using the *fitlm* command in MATLAB R2017b:

$$ln\frac{C_{eff}}{C_{inf}} = B_0 + B_1(k_{bio} \times EBCT) + B_2(EBCT) + B_3\left(C_{inf}\right) + B_4\left(\frac{BV}{C_{inf}}\right) + B_5(logD_{ow}) + B_6(Z) + B_7 DOC_{inf} + B_8 T_m \quad (2-4)$$

where: B_i are the regression coefficients; C_{eff} is the effluent CEC concentration (ng/L); C_{inf} is the influent CEC concentration (ng/L); k_{bio} is the biodegradation rate constant (min^{-1}), k_{bio} values for both low (3-10 ^oC) and high (20-28 ^oC) water temperature were computed for each CEC and used in the analysis; *EBCT* is empty bed contact time (min); $k_{bio} \times EBCT$ for each CEC was used to represent CEC removal via biodegradation; *BV* is throughput (as bed volumes); BV/C_{inf} was used to represent the fraction of occupied sorption sites in GAC-sand biofilters (a detailed explanation is provided in the Appendix B); $logD_{ow}$ is the CECs octanol-water partitioning coefficient (i.e., the $logK_{ow}$ value corrected for the fraction of the compound in the neutral form at pH 8.5, Table 2-2); *Z* is CEC charge (at pH 8.5, Table 2-2); DOC_{inf} is the influent DOC concentration (Fig. B-2; mg/L); T_m is the influent water temperature (^oC).

The p-value for each regression coefficient was also calculated. Parameters with no significant effect on CECs removal at 95% confidence level were excluded from the final empirical relationship.

2.3 Results

2.3.1 Batch sorption experiments

Batch sorption experiments were performed to determine the sorption capacity of the GAC media. No significant reduction of any of the selected CECs was observed in control bottles without GAC after 5 days of incubation in buffered DI water (pH 8.5) at room temperature (data not shown), suggesting that the selected compounds have low or negligible hydrolysis rates. The data from the batch sorption tests in buffered ultrapure water (pH 8.5) at room temperature (23 ^{o}C) were fit to the Langmuir isotherm model and the isotherm parameters are summarized in Table 2-3. Similar results were obtained for experiments in which the water also contained 5.0 ppm as C of Mississippi River NOM (Table B-3).

Table 2- 3. Langmuir isotherm parameters (mean ± standard error) for the selected CECs obtained by fitting the batch test results for the following conditions: room temperature, 0.1 M carbonate buffered ultrapure water (pH 8.5).

	K_l (L/mg_CECs)	q_m (mg_CECs/mg_AC)	K_d (L/mg_AC)	R^2
Atenolol	160.3±38.3	0.027±0.001	4.33	0.84
Atrazine	139.9±39.7	0.033±0.001	4.62	0.84
Carbamazepine	189.7±51.7	0.033±0.001	6.26	0.82
Fluoxetine	26.0±6.6	0.110±0.006	2.86	0.91
Gemfibrozil	133.7±47.2	0.050±0.003	6.68	0.87
Metolachlor	21.9±7.4	0.070±0.005	1.54	0.79
Sulfamethoxazole	70.1±25.7	0.011±0.001	0.77	0.71
TCEP	312.2±99.4	0.015±0.001	4.68	0.60

2.3.2 CECs removal performance of the anthracite-sand biofilters

The selected CECs have negligible hydrolysis rates, hence CECs removal in the anthracite-sand biofilters is attributed solely due to biodegradation. The removal efficiencies of all selected CECs in the anthracite-sand biofilters for all sampling periods are shown in Fig. 2-2.

Among the eight CECs, fluoxetine had the highest removal over the entire course of the experiment, ranging from 13.3 to 62.8% at low temperature (3-10 ^{o}C) and from 49.2 to 97.3% at high temperature (20-28 ^{o}C). These results suggest that conventional biologically-active anthracite-sand filters are an effective method to remove fluoxetine during drinking water treatment. Low to moderate biodegradation of metolachlor, atenolol, gemfibrozil, and sulfamethoxazole (10 to 50%) was also observed in the anthracite-sand biofilters, but only at high temperature (20-28 ^{o}C). No significant removal of atrazine, carbamazepine, and TCEP was observed in the anthracite-sand biofilters over the entire course of the experiment, indicating that these compounds are recalcitrant to biodegradation in this system.

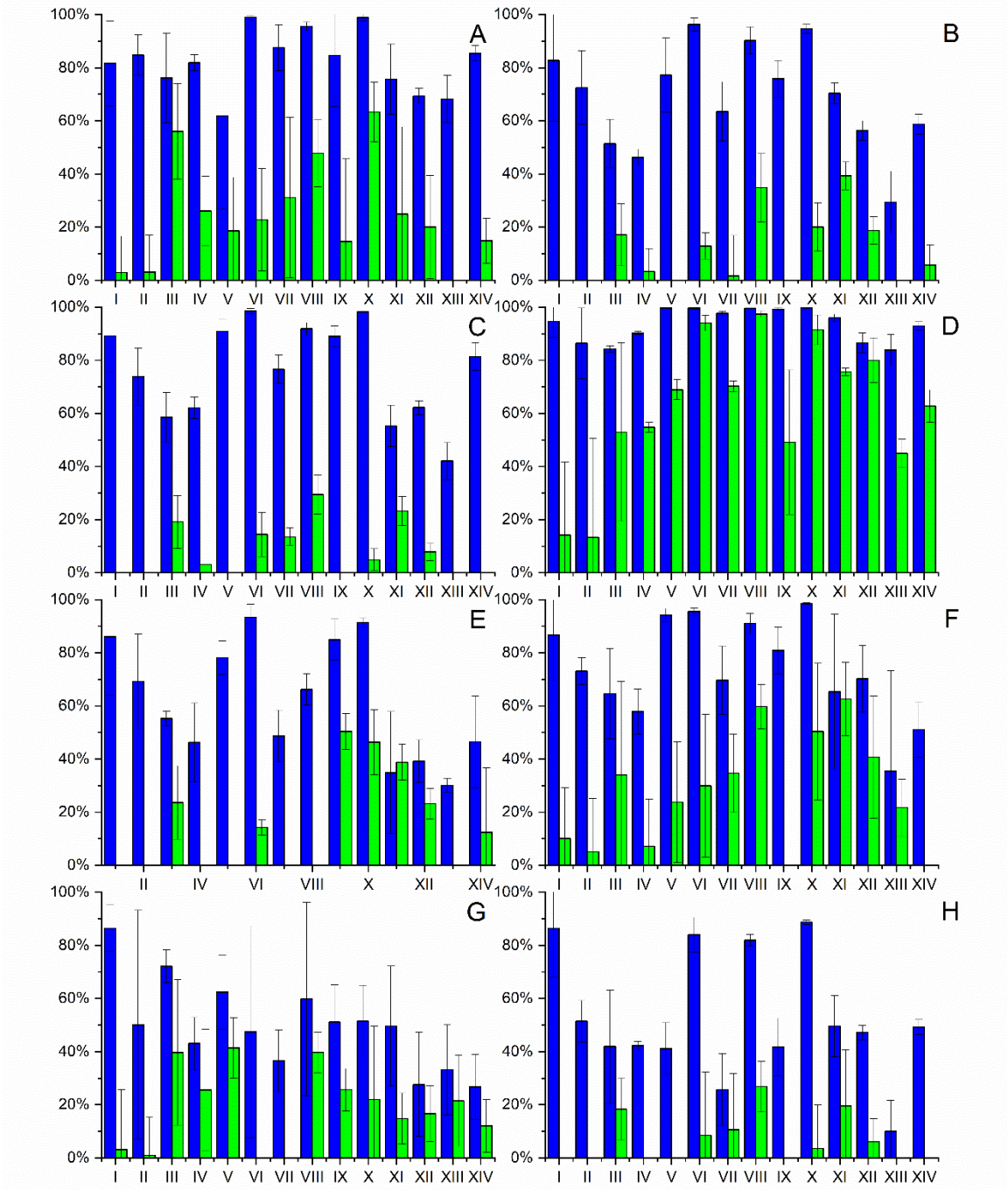

Figure 2- 2. Mean removal efficiencies (in %) of atenolol (**A**), atrazine (**B**), carbamazepine (**C**), fluoxetine (**D**), gemfibrozil (**E**), metolachlor (**F**), sulfamethoxazole (**G**), and TCEP (**H**) in the pilot-scale GAC-sand (blue) and anthracite-sand (green) biofilters over 14 sampling periods.

The k_{bio} values are summarized in Table 2-4 for both high temperature and low temperature conditions. Fluoxetine was the only CEC with a k_{bio} value significantly greater than zero ($p=2.47\times10^{-3}$) for the low temperature condition (3-10 ^{o}C). As expected, at high temperature (20-28 ^{o}C) the k_{bio} values of the CECs (0.05±0.03 to 0.93±0.15 min^{-1}) were greater than those at low temperature (Table 2-4).

Table 2- 4. The biodegradation rate constant (kbio; mean ± standard error) for each selected CEC at two temperature conditions determined from the anthracite-sand biofilter CEC removal results.

	Biodegradation rate constant k_{bio} (min^{-1})	
	Low Temperature	**High Temperature**
Atenolol	0.05±0.03	0.20±0.04*
Atrazine	-0.02±0.02	0.09±0.04*
Carbamazepine	-0.03±0.02	0.05±0.03
Fluoxetine	0.22±0.06*	0.93±0.15*
Gemfibrozil	-0.02±0.02	0.11±0.04*
Metolachlor	0.02±0.02	0.29±0.06*
Sulfamethoxazole	0.07±0.04	0.11±0.03*
TCEP	-0.03±0.02	0.06±0.03

Note: High temperatures ranged from 20-28 ^{o}C; Low temperatures ranged from 3 to 10 ^{o}C.
* indicates the mean value is significantly greater than zero at the 95% confidence level.

No significant correlation (p=0.617) was observed between BIOWIN6 predictions and k_{bio}, suggesting that this structure-activity-based model is not a useful predictor of CECs biodegradation in biofilters. Preliminary work using ion trap mass spectrometry was performed to detect possible biodegradation products, but the results were inconclusive partly due to a high background signal attributed to matrix interferences.

2.3.3 CECs removal performance of the GAC-sand biofilters

The eight selected CECs in the GAC-sand biofilters were moderately to effectively removed over the entire course of the experiment, with mean (± standard deviation) removals ranging from 49.1±27.7 % to 94.4±7.0 %. The removal efficiencies of three selected CECs (carbamazepine, metolachlor and fluoxetine) in the GAC-sand biofilters throughout all sampling periods are shown in Fig. 2-2. For all selected CECs, GAC-sand biofilters provided superior removal compared to anthracite-sand biofilters, due to the dual function of adsorption and biodegradation.

The biofilters were fed with water from full-scale recarbonation chambers, which contained 2.95 to 8.20 ppm DOC during the course of the study (Fig. B-2). Removal of DOC in the GAC-sand biofilters was initially 15 to 22 % but stabilized at 1 to 5% after 4 months of operation (Fig. B-2), suggesting that the GAC filter media was essentially saturated with DOC. Despite DOC breakthrough, effective removal of CECs was observed in the GAC-sand biofilters even after 19 months of operation.

The fraction of sorption sites occupied (q_c/q_e) was computed for each CEC in the GAC-sand biofilters throughout the experiment, where qe was computed assuming an aqueous concentration of 100 ng/L (Fig. 2-3). Among all selected CECs, sulfamethoxazole had the highest q_c/q_e at the end of the experiment (0.286) and gemfibrozil had the lowest (0.025). Assuming carbon regeneration is required when 90% of the GAC media is saturated with a target CEC, the estimated frequency of regeneration ranged from 1232 to 14244 days (3.4 to 39.0 years). Both adsorption and biodegradation need to be considered to explain the differences of q_c/q_e among the CECs. In general, CECs with higher K_d have lower end-of-experiment q_c/q_e due to higher adsorption capacities. Some exceptions are noted for biodegradable CECs. For example, even though fluoxetine had a lower K_d than that for TCEP (2.86 and 4.68, respectively), the end-of-experiment q_c/q_e value of fluoxetine was still lower than TCEP (0.048 and 0.059, respectively) due to extensive biodegradation of fluoxetine.

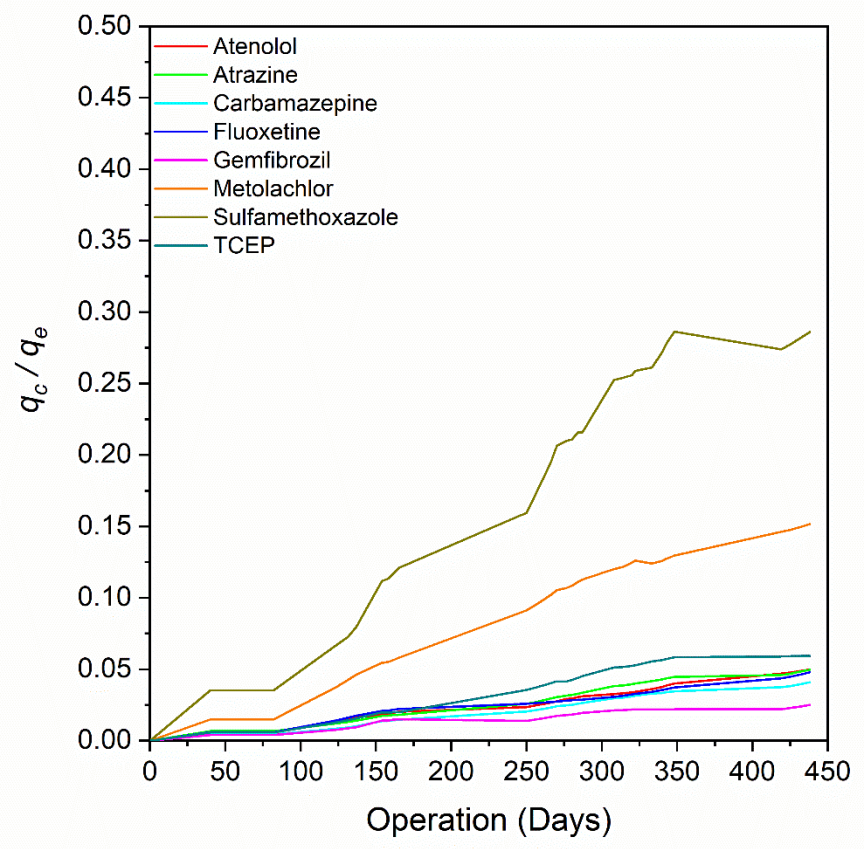

Figure 2-3. Ratio of the amount of CEC sorbed to the GAC media (q_c) to the estimated sorption capacity assuming an aqueous CEC concentration of 100 ng/L (q_e) for all selected CECs over time.

Effect of hydraulic loading rate. CEC removal in the GAC-sand biofilters was significantly greater ($p=8.33\times10^{-9}$) at the low loading rate (Fig. 2-4). The mean removal efficiency decreased by 16.5 % when the loading rate increased from 2 to 4 gpm/ft^2 (4.88 to 9.76 m/h, EBCT: 8.4 min to 4.2 min).

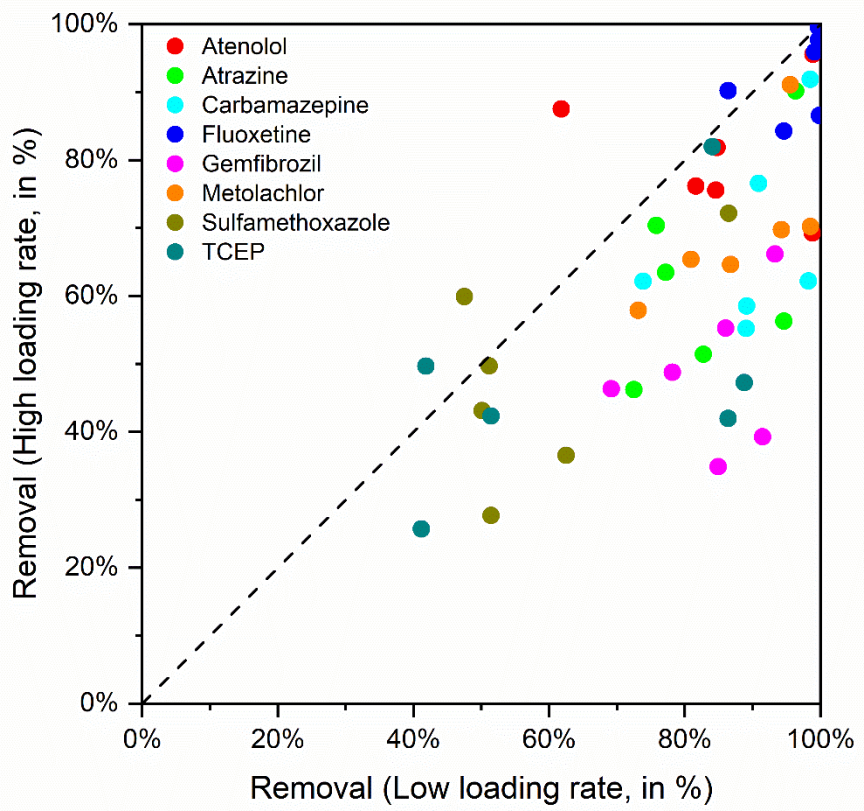

Figure 2-4. Effect of hydraulic loading rate on CEC removal in GAC-sand biofilters. The removals for the high loading rate (4 gpm/ft^2 or 9.76 m/h) periods of III, IV, VI, IIX, XI and XII were compared to the removals for the low loading rate (2 gpm/ft^2 or 4.88 m/h) periods of I, II, V, VI, IX, and X. The dashed line represents the 1:1 ratio.

Effect of throughput. To investigate the effect of throughput on CEC removal performance, the GAC-sand biofilters were operated under the same conditions for periods I and XIII and periods II and XIV, which occurred approximately 100,000 bed volumes of throughput apart (Table 2-1). The biofilter performance for these periods is compared in Fig. 2-5. Significant reductions of CECs removal (p: 3.92×10^{-8} to 2.87×10^{-2}) in GAC-sand biofilters for all selected CECs were observed after 100,000 bed volumes of throughput when the influent CECs concentration was low (100-200 ng/L, Fig. 2-5-A). CECs removal

efficiencies declined by 10.8% (fluoxetine) to 76.4% (TCEP). No significant reduction in CEC removal performance (p: 0.056 to 0.430) for most selected CECs (except metolachlor) was observed when the influent CECs concentration was high (1000-3000 ng/L, Fig. 2-5-B).

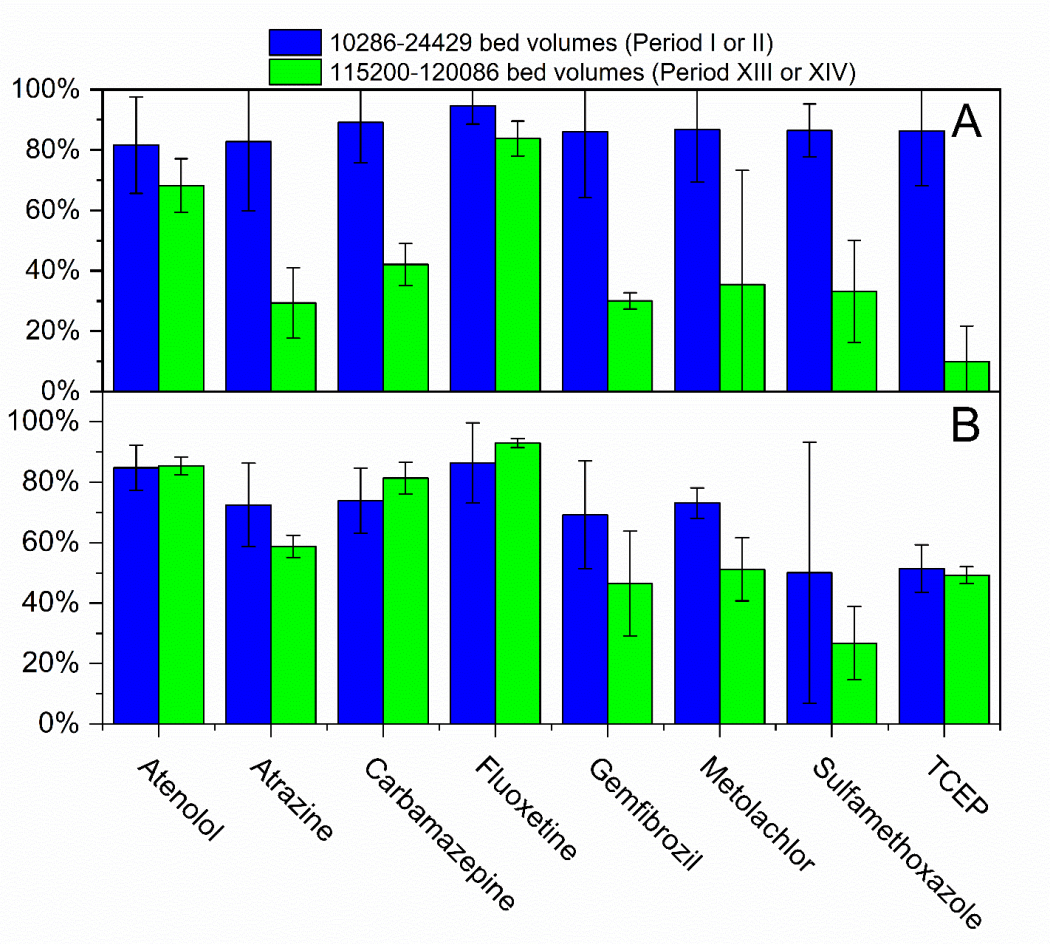

Figure 2-5. Effect of throughput on CEC removal in the GAC-sand biofilters under conditions of low temperature (3-10^{o}C) and low loading rate (2 gpm/ft^{2} or 4.88 m/h). A: Period I vs. Period XIII with low influent CECs concentrations (100-200 ng/L); B: Period II vs. Period XIV with high influent CECs concentrations (1000-3000 ng/L). Error bars represent one standard deviation from the mean values.

2.3.4 Multiple linear regression analysis to determine the most important factors affecting CECs removal in the GAC-sand biofilters

Three factors had coefficients that were not statistically significant in the multiple linear regression: influent CECs concentration (B_3), hydrophobicity ($\log D_{ow}$ at pH 8.5, B_5), and temperature (B_8) (p=0.795, 0.094, and 0.226, respectively). After excluding insignificant parameters, the final empirical relationship is shown in Eq. 2-5. (N=220, $p=3.15\times10^{-49}$, $R^2=0.668$, $R^2_{adj}=0.660$),

$$ln\frac{C_{eff}}{C_{inf}} = 1.270 + (-0.482 \pm 0.043)k_{bio} \times EBCT + (-0.280 \pm 0.045)EBCT + (0.0015 \pm 0.0002)\frac{BV}{C_{inf}} + (-0.537 \pm 0.094)\,Z + (-0.325 \pm 0.047)DOC_{\ inf} \qquad (2-5)$$

and comparison of observed C_{eff}/C_{inf} values and C_{eff}/C_{inf} values computed using Eq. 2-5 is shown in Fig. 2-6. The adjusted correlation coefficient of 0.660 indicates that 66.0 % of the variation in C_{eff}/C_{inf} for the pilot-scale GAC-sand biofilters can be explained by this empirical relationship.

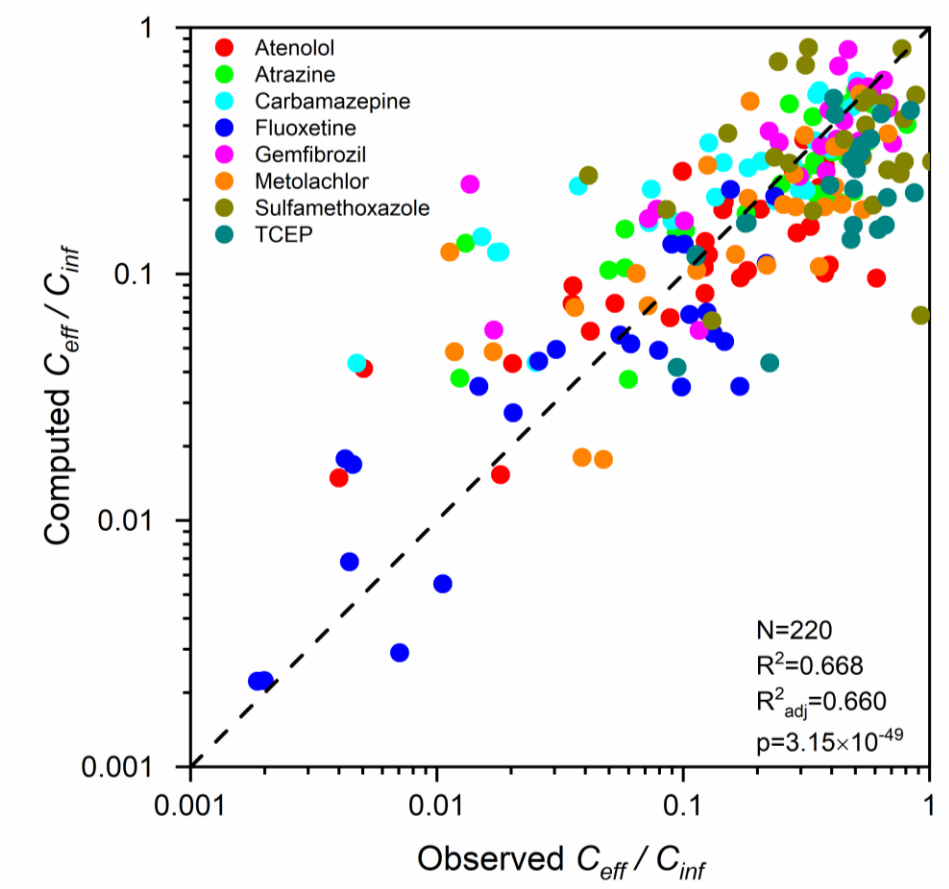

Figure 2-6. Comparison of values of the effluent to influent CEC concentration ratios (C_{eff}/C_{inf}) computed using the empirical relationship developed in the present study to the observed C_{eff}/C_{inf} values. The dashed line represents the 1:1 ratio. The adjusted correlation coefficient (R^2_{adj}) equals 0.660, meaning 66.0% of the variation in C_{eff}/C_{inf} can be explained by using this relationship.

The coefficients for $k_{bio} \times EBCT$, *EBCT*, and BV/C_{inf} (-0.482±0.043, -0.280±0.045, 0.0015±0.0002, respectively) indicated that, as expected, increasing biodegradation rate and *EBCT* results in increased CEC removal (i.e., reduced C_{eff}/C_{inf}) while increasing the fraction of occupied sorption sites (represented as BV/C_{inf}) decreased CEC removal. All three factors were significant and should be considered when modeling CEC removal performance in GAC biofilters. The coefficients for CEC charge (*Z* at pH 8.5) and influent

DOC concentration (-0.537±0.094 and -0.325±0.047, respectively) suggest that CECs are more readily removed in GAC biofilters when the compounds are positively charged and, surprisingly, when the influent DOC concentration increases.

Effluent CEC concentrations from simulations using AdDesignS software (Fig. B-3) were in poor agreement with experimental observations. For carbamazepine, gemfibrozil, and sulfamethoxazole, the simulated effluent concentrations were consistently lower than the observed values. The simulated effluent concentrations of metolachlor were constantly higher than the observed values. For atenolol, atrazine, fluoxetine and TCEP, the simulated effluent concentrations were lower than the observed values for the first 200 days of operation then higher thereafter.

2.4 Discussion

This study provided a comprehensive investigation of CECs removal in GAC-sand biofilters used for drinking water treatment. A unique aspect of this study was the operation of GAC-sand and anthracite-sand biofilters in parallel under actual drinking water treatment conditions with continuous CECs dosing for an extended period of time (> 1 year). By comparing the performance of GAC-sand and anthracite-sand biofilters operated in parallel, the relative roles of adsorption and biodegradation were able to be estimated. Variation in operating and environmental conditions permitted evaluation of the effect of contaminant concentration, filtration rate or EBCT, and water temperature on CECs removal. These findings should be helpful to drinking water treatment facilities in predicting possible CEC removal and in optimizing the biofiltration process to maximize CECs removal.

The contribution of biodegradation to CECs removal in the biofilters varied significantly from compound to compound. Atrazine and carbamazepine were recalcitrant to biodegradation in the pilot-scale biofilters (mean removal efficiencies<10%). These findings are consistent with previous reports in the literature [5, 21, 23]. The consistency of observations between several studies suggests that these three CECs are not treatable using conventional biofiltration, regardless of filter design and influent water quality. The k_{bio} value of fluoxetine was significantly greater than zero throughout the experiment, suggesting conventional biofilters can be an effective method to remove fluoxetine in

drinking water treatment. Other selected CECs had k_{bio} values that were not significantly greater than zero (p: 0.076 to 0.495) at low temperature (3-10 ^{o}C), suggesting conventional biofiltration could be less effective in removing CECs from cold water, even for some relatively biodegradable compounds (i.e., metolachlor and atenolol). Our k_{bio} values at high temperature (20-28 ^{o}C) were similar to those reported by Zearley and Summers (2012), but there were some differences for individual CECs. For example, the mean k_{bio} for metolachlor of 0.29±0.06 min^{-1} (20-28 ^{o}C) was about 32 times greater than the previously reported value (0.009 min^{-1} at 20 ^{o}C). This difference in biodegradation rate is possibly due to differences in the amount of active biomass in the biofilters and feed water quality between the two studies. The large difference in biodegradation rate also suggests that k_{bio} values are system-specific and methods to standardize biodegradation rate constant in biofilters need to be developed. Other than temperature, pH, ozonation and superficial velocity can also significantly affect CECs removal by biofiltration [24, 34, 35].

Understanding CECs biodegradation pathways is critical for estimating potential reductions in overall toxicity and may help drinking water treatment facilities optimize biofilter performance. Microbial hydrolysis of the amide bond on atenolol was observed in biofilters and wastewater bioreactors [35, 36]. Metolachlor is normally considered persistent in natural waters, but microbial dechlorination has been observed in soil [37, 38]. Biodegradation of fluoxetine was observed in an activated sludge bioreactor, likely due to oxidation or hydroxylation followed by dealkylation [39]. The results of biodegradation product identification in the present study were inconclusive, likely because of the low

concentrations of CECs in the feed and high background NOM concentrations which may interfere with product detection. Additional research is needed to identify the dominant CECs biodegradation products and elucidate the relevant pathways in biofilters.

The relative roles of sorption and biodegradation in the GAC-sand biofilters were estimated by comparing the CECs removal results for the GAC-sand filters with those in the anthracite-sand biofilters. This study is unique in that other investigations concerning CECs removal from drinking water supplies involved non-sorptive media such as sand or anthracite [5, 21–23] or "spent" GAC [24]. It can be argued that the validity of this approach to determining the relative roles of sorption and biodegradation requires that the microbial activity levels and compositions of the communities in the GAC biofilters and anthracite biofilters are similar, which may not always be the case [15]. An investigation of the microbial communities in the media from the biofilter columns in the present study is the subject of ongoing research.

For all selected CECs except fluoxetine, GAC-sand biofilters provided at least double the removal in the anthracite-sand biofilters (Fig. 2-2). Assuming that the rate of biodegradation was the same in both types of filters, this result suggests that adsorption was the dominant CECs removal mechanism in GAC-sand biofilters over the first 15 months of operation with initially virgin GAC. The assumption of similar biodegradation rates seems reasonable because ATP concentrations on the filter media were similar for the GAC-sand and anthracite-sand biofilters (Fig. B-1). The abundance and activity of specific

CEC-degrading bacteria would be needed to confirm the validity of this assumption, but this would be challenging to determine and was beyond the scope of this study.

The observed effects of hydraulic loading rate on GAC-sand biofilter performance were expected because low loading rate provides greater contact time for both adsorption and biodegradation. Similar results also were found in other studies at a smaller scale, involving rapid small-scale column tests (RSSCT) or bench-scale GAC filters [33, 40, 41].

The difference between the fraction of occupied GAC sorption site (q_c/q_e) values for biodegradable and non-biodegradable CECs suggests that biodegradation can extend filter bed life, either by reducing the initial amount of sorption or by enhancing desorption and regeneration of occupied sorption sites, which is consistent with observations of geosmin removal in GAC biofilters [10]. The effect of throughput on CECs removal is explained by considering the usage of sorption sites. For non-biodegradable CECs, a greater decrease in removal efficiency was observed for CECs with higher q_c/q_e (for example: TCEP vs. carbamazepine). Biodegradable CECs, such as fluoxetine and atenolol, were less impacted by throughput because less sorption sites were occupied over time due to the significant role of biodegradation in removal. The lack of an effect of throughput at high influent CECs concentrations (1000-3000 ng/L) is explained by the increase in driving force for sorption when the influent is suddenly switched from low to high concentration, so that carbon saturated at the lower concentration is able to sorb more CECs.

A multiple linear regression analysis was performed to investigate factors that significantly affected CEC removal in the GAC-sand biofilters. Unlike some previous studies [22, 42], the D_{ow} of the compounds was not a significant factor in determining CECs removal. It is important to note that many of the CECs in this study have polar and/or charged functional groups which may also affect CECs sorption onto GAC [31, 43]. Our results suggested that positively charged CECs were more sorptive than neutral or negatively-charged compounds. Typical point of zero charge pH values (pH_{pzc}) for activated carbon are between 3.4 to 7 [44–46]. Thus, the virgin GAC media surface is expected to be negatively charged at pH 8.5 (i.e., the influent water pH). The enhanced adsorption of positively-charged CECs was likely due to electrostatic attraction [31, 47].

One unexpected result from the linear regression analysis was the negative coefficient for the DOC term, suggesting that increasing DOC concentration was beneficial for CECs removal. DOC in filter influent is often associated with a reduction in contaminant removal efficiency and earlier breakthrough due to competition for adsorption sites and pore blocking [3, 48]. Considering DOC is mostly negatively charged in natural water, increasing influent DOC concentration may lead to a stronger electrostatic attraction for some compounds once DOC absorbs to the GAC. For example, Ridder et al. (2011) reported stronger electrostatic attraction of pharmaceuticals on GAC preloaded with NOM compared to fresh GAC [31]. In addition, as co-metabolism is the likely biodegradation mechanism given the low concentrations of these compounds, higher DOC levels also

should provide more labile carbon for increasing biomass and bioactivity levels in the filter media.

Developing models for predicting CEC removal in GAC-sand biofilters is challenging given the complexity of the process. First, the possible removal mechanisms include sorption, biodegradation, and chemical degradation (e.g., hydrolysis). Second, the performance is expected to be highly dynamic with short-term variation in biodegradation rates due to the effects of backwashing on attached biomass levels [49] as well as moderate to longer term changes in sorption effectiveness due to accumulation of CECs and NOM on the carbon surface and within the pores [22, 50]. Further, the use of GAC as a filter media means that particles and flocs can accumulate on the carbon and possibly block pores [51] or even provide more sites for sorption of CECs [52, 53]. We explored the use of AdDesignS for predicting the CECs removal performance of the GAC-sand biofilters in this investigation. The AdDesignS program incorporates adsorption capacity, adsorption kinetics and NOM fouling, but does not account for biodegradation and water quality variations (i.e., influent DOC concentration). The agreement between the model predictions and experimental data was poor in some cases, especially for compounds that were biodegradable [10]. Therefore, an empirical relationship was developed via multiple linear regression that incorporated all potential factors that may affect CECs removal in GAC-sand biofilters, including biodegradation rate, adsorption capacity, throughput, EBCT, water quality variation (influent DOC concentration) and CEC chemical properties (charge). Unfortunately, the problem with such empirical relationships is that they tend to

be site specific. Additional research is needed to determine how well the empirical relationship developed in this study applies to other systems.

2.5 Acknowledgements

Financial support for this work was provided by the City of Minneapolis. This work was carried out in part using LC-MS/MS instruments at Masonic Cancer Center at University of Minnesota, Twin Cities. We thank Annika Bankston, George Kraynick and Dr. Li Zhang of Minneapolis Water Treatment and Distribution Services for helping coordinate this study and providing technical input. We thank Lucian Osuji, Christopher Rydell, Andy Weyer, Joe Kroening, Troy Rosenthal and Eric Raway for helping with sample collection and water quality analysis. We also thank Dr. Peter Villalta, Xun Ming and Dr. Jill Kerrigan for help in the development of our SPE and LC-MS/MS methods.

Chapter 3

Investigation of the microbiome in the pilot-scale biofilters and the impact of biofilters on the microbiome in filter effluent.

Bacterial communities in biofilters can be beneficial through biodegradation of contaminants but also pose potential risks by harboring and releasing detrimental microbes into water distribution systems. This study investigated the variations in the bacterial communities in biofilters and its relationship with the pre-/post-filtration water microbiome by operating a pilot-scale biofiltration system for over 18 months. The bacterial abundance and community composition on filter media samples (n=512) at four different depths (top layer, 0.15m, 0.3m, and 0.45m), filter influent samples (n=39) and filter effluent samples (n=104) were documented by analyzing 16S rRNA gene using real-time quantitative polymerase chain reaction and Illumina HiSeq high-throughput sequencing. The bacterial abundance in the biofilters was relatively stable over time. But the community composition exhibited temporal variations and varied with media type (GAC vs. anthracite), backwash strategy (chloraminated vs. non-chloraminated), and bed depth. The biofilters effectively removed biomass (~70%) from the water but only marginally impacted the bacterial community composition in the filter effluent through collecting and shedding. This study provided a better understanding of the microbiome in biofilters from both ecological and engineering perspectives and possible approaches for managing biofilter microbiomes.

3.1 Introduction

Filtration using biologically active, granular media has been widely used in drinking water treatment since 1800s [12]. The bacterial community on filter media can be beneficial in removing of natural organic matter [8, 9], taste-and-odor compounds [10], and other emerging contaminants [11]. But as a part of the source-to-tap microbiome, it may also pose potential risks in terms of public perception by harboring and releasing detrimental microbes into drinking water distribution systems [14, 15]. To improve the performance and manage the potential risks, it is essential to understand the bacterial community in biofilters and how it shapes the microbiome in drinking water.

There are many factors affecting the variations and dynamics in bacterial communities in both natural and engineered ecosystems. Most bacterial communities exhibit spatial and temporal heterogeneities due to the changes in environmental factors (e.g., temperature, pH, and carbon availability) as well as other stochastic events like growth, migration, and predation [54–56]. Microorganism-specific selection of lifestyle (e.g., biofilm vs. planktonic) is also critical in determining the bacterial community assembly, especially in a complex ecosystem involving multiple types of living environments [57]. Highly reactive chemicals, such as disinfectants (e.g., chloramine) and contaminants of emerging concerns (CECs; including micropollutants like pesticides and antibiotics), are detected in various environments and also significantly impact bacterial communities even in trace concentrations [58–61].

It is challenging to investigate the bacterial community in biofilters, because it is a complex ecosystem influenced by both environmental factors and engineering designs. There are a few studies concerning the bacterial community in biofilters. Bacteria like *Alphaproteobacteria*, *Gamaproteobacteria*, *Nitrospira* (class level) were consistently detected in full-scale biofilters [13, 14]. The bacterial community on the filter media exhibited temporal variation and may help shaping the microbiome in water distribution systems [13–15]. Despite the previous works, little is known about how the bacterial community in biofilters responds to the water quality changes, reactive chemicals exposure, and different filter operational conditions (e.g., backwash strategies) and how the community on filter media interact with the pre-/post-filtration water microbiome.

In this study, the microbiome in biofilters was investigated by operating eight pilot-scale biofilters in parallel for over 18 months at a full-scale water treatment plant. The bacterial abundance and high-resolution community profile at genus level were well documented using latest developed analytical techniques, including real-time quantitative polymerase chain reaction (qPCR) and Illumina HiSeq high-throughput sequencing. The temporal and spatial (at different bed depths) variations and the effects of water quality change, filter media type (granular activated carbon (GAC)-sand vs. anthracite-sand), CECs addition and backwash strategy (using chloraminated water vs. non-chloraminated water) on the bacterial community in biofilters were evaluated. Finally, the impact of biofilters on the microbiome in filter effluent was investigated.

3.2 Methods

3.2.1 Pilot-scale biofiltration system design and operation

A pilot-scale biofiltration system was constructed at the Minneapolis Water Treatment and Distribution services (MWTDS) in Fridley, Minnesota, USA, as described in Chapter 2. Briefly, water from the full-scale recarbonation chambers was used as the feed water for the pilot plant (Fig. 2-1). The water was adjusted to a pH of 8.5 using hydrochloric acid, coagulated using 2 mg/L (as Fe) ferric chloride, and then settled for 1 to 2 hours before entering the biofilter columns. Eight biofilter columns were operated in parallel. Five columns (column #1-3, 7 and 8) were GAC-sand biofilters and three columns (column #4-6) were anthracite-sand biofilters.

The pilot-scale biofiltration system was operated for 18 months from 07/15/2015 to 01/04/2017. The filters were operated at hydraulic loading rates of 4.9 or 9.8 m/h according the operational schedule (Table 2-1). The columns were backwashed every 4 to 6 days, using 36.8 L/min of air for 5 minutes followed by 10 minutes of water wash at 36.66 m/h to achieve a 30% bed expansion. Two backwash strategies were used: Chloraminated backwash, which used finished water from MWTDS containing 2-4 mg/L (as Cl_2) monochloramine, was applied from 07/15/2015 to 08/15/2016; Non-chloraminated backwash, which used the filter effluent directly from the pilot plant without chloramine, was applied from 08/16/2016 to 01/04/2017. Environmentally relevant concentrations of

CECs (0.1-3 µg/L) were continuously dosed into the filter influent of six columns (column #1-6), as described in Chapter 2. Eight CECs that have been detected frequently in finished drinking water [1] were selected for this study, including herbicides (atrazine and metolachlor), pharmaceuticals and antibiotics (atenolol, carbamazepine, fluoxetine, gemfibrozil, and sulfamethoxazole), and an industrial chemical (tris (2-chloroethly) phosphate).

The water temperature was measured daily. The concentrations of DOC and total ammonia in the filter influent samples were determined weekly according to EPA method 415.3 and 350.1, respectively.

3.2.2 Sample collection and DNA extraction

Water samples and media samples from the biofilters were collected for analysis of bacterial abundance and community composition. Filter influent and effluent samples (1 L) from all biofilter columns were collected monthly in pre-sterilized polypropylene bottles. Filter media samples (approximately 0.5 g as wet weight) were collected in sterile microcentrifuge tubes right after water sample collection from four different depths (top, 0.15 m, 0.30 m and 0.45 m) in all biofilter columns. All samples were transported on ice to the laboratory. Upon arrival, water samples were filtered through pre-sterilized nitrocellulose filters (diameter = 47mm; pore size = 0.2 µm). The nitrocellulose filters and media samples were then submerged in 0.6 mL lysis buffer (120 mM sodium phosphate,

5% SDS, pH 8.0), subjected to three consecutive freeze-thaw cycles and incubated at 70 ^{o}C for 90 minutes. Then, genomic DNA in each sample was extracted using the FastDNA SPIN Kit for Soil (MP Biomedicals, Solon, OH). Purified DNA was stored at -20°C until future use. After extraction, the media samples in the microcentrifuge tubes were removed, dried at 105 ^{o}C, and then weighed to obtain the dry weights of the media for data normalization purposes. In total, 143 water samples and 512 media samples were collected in this investigation.

3.2.3 Real-time quantitative polymerase chain reaction (qPCR)

The 16S rRNA gene for Bacteria that was used to quantify total bacterial biomass in water samples and media samples using real-time quantitative qPCR [62]. Standards for qPCR were prepared from serial dilution of synthetic double-stranded DNA fragments (gBlocks) to develop calibration curves for quantification of gene copies in each sample. The censoring limit (CL) was defined as the lowest standard of a given qPCR assay that amplified (1100 copies per reaction), normalized by the volume (L) of the water sample or the bed volume (cm^3) of the media sample used for DNA extraction. The detailed information of the gene targets and respective primer sequences and gBlock standards is shown in Table C-1. A paired, two-sample t-test was used to determine if there was a significant difference in gene concentrations between GAC and anthracite media samples. A Pearson's correlation test was conducted to determine if a statistically significant correlation existed between two results (e.g., gene concentration and water temperature).

3.2.4 High-throughput 16S rRNA gene sequencing

A total of 319 samples were selected for sequencing analysis, including all water samples, all media samples collected from the top layer of the biofilters, and the media samples collected from four different depths on 07/28/2016 and 09/28/2016. The DNA extract from each sample was amplified using PCR, purified, quantified, and then pooled by equal mass to create a library for sequencing, as previously described [13]. Next generation high-throughput sequencing of PCR-amplified 16S rRNA gene fragments (V3 region) was performed on the Illumina HiSeq platform at the University of Minnesota Genomics Center (UMGC) using primers 341F and 534R [63]. Unprocessed sequence reads have been uploaded to Sequence Read Archive (accession PRJNA515121).

Sequence reads were demultiplexed, trimmed, and filtered using QIIME2 (version 2018.2). Amplicon sequence variants (ASVs) were determined using 'DADA2' [64, 65] and then assigned consensus taxonomy using SILVA rRNA database (release 128) [66]. A phylogenetic tree was generated using 'FastTree2' and the rooted using midpoint [67].

Both alpha (Shannon index) and beta (weighted UniFrac distance metrics) diversity were calculated using QIIME2 (pipeline: core-metrics-phylpenetic) after sequence libraries were randomly trimmed down to 100,000 sequences per profile. Paired two-sample t-test was used to determine if there was a significant difference in Shannon index values between GAC and anthracite media samples. Principal coordinates analyses (PCoA) plots were

made to visualize the weighted UniFrac distance metrics. “Core-satellite” model was used to determine the core ASVs based on their occupancy (i.e., the proportion of the samples covered) and the mean relative abundance [68, 69]. Multiple correlation analyses were conducted to determine the correlation between the relative abundance of individual ASV and water quality parameters (e.g., temperature, influent DOC and ammonia concentrations), using *corr* command in MATLAB 2018a, and the p-values were adjusted using Bonferroni correction [70]. Differential abundance analyses were performed to determine the identities of ASVs which relative abundance were significantly different between sample groups, using ‘DESeq2’ package in R software, as previously described [60]. Permutational multivariate analysis of variance (PERMANOVA) tests (999 permutations) were conducted using the *adonis* function in R software to determine the effects of water quality parameters on the bacterial community. *Post hoc* analyses of pair-wise comparisons were performed using ‘pairwiseAdonis’ package in R. The dispersions of the bacterial community within a sample group were tested using the *betadisper* and *permutest* functions in ‘vegan’ (999 permutations), and *post hoc* analyses were performed using Wilcoxon tests. For all *post hoc* tests, p-values were adjusted using Benjamini-Hochberg (BH) procedure [71].

3.3 Results

3.3.1 Biofilter influent water quality

Influent water temperature, total ammonia concentration, and DOC concentration were regularly monitored over the entire course of the experiment (Fig. C-1). Temperature ranged from 3 to 28 ^{o}C, with a mean of 12.6 ^{o}C. DOC concentrations ranged from 2.9 to 8.2 mg C/L, with a mean of 4.8 mg C/L. Total ammonia concentration ranged from below the detection limit (0.015 mg N/L) to 0.13 mg N/L, with a mean of 0.05 mg N/L.

3.3.2 Quantification of total bacterial biomass

The 16S rRNA gene concentration in the media samples collected from the top layer fo the GAC-sand and anthracite-sand biofilters ranged from 8.4 ± 0.2 to 9.8 ± 0.1 and 8.0 ± 0.2 to 9.7 ± 0.0 log copies/cm^3 bed volume, respectively (Fig. 3-1). No significant difference (P=0.272) in 16S rRNA gene concentrations was observed between GAC and anthracite media. Also, the 16S rRNA gene concentrations were relatively stable over time and did not correlate with water temperature. The mean concentrations of 16S rRNA genes in the media samples decreased with depth in the biofilters (Fig. C-2), with a mean reduction of 0.6 ± 0.4 and 0.6 ± 0.7 log copies/cm^3 bed volume for every 0.15 m of depth increase in the GAC-sand and anthracite-sand biofilters, respectively. The 16S rRNA gene concentrations in the filter influent ranged from 6.9 ± 0.3 to 8.8 ± 0.0 log copies/L (Fig. C-

3). The GAC-sand and anthracite-sand biofilters effectively removed biomass (i.e., 16S rRNA genes) from the filter influent, providing 64.9±30.2% and 70.0±34.6% removal, respectively, over the course of the experiment (Fig. C-3).

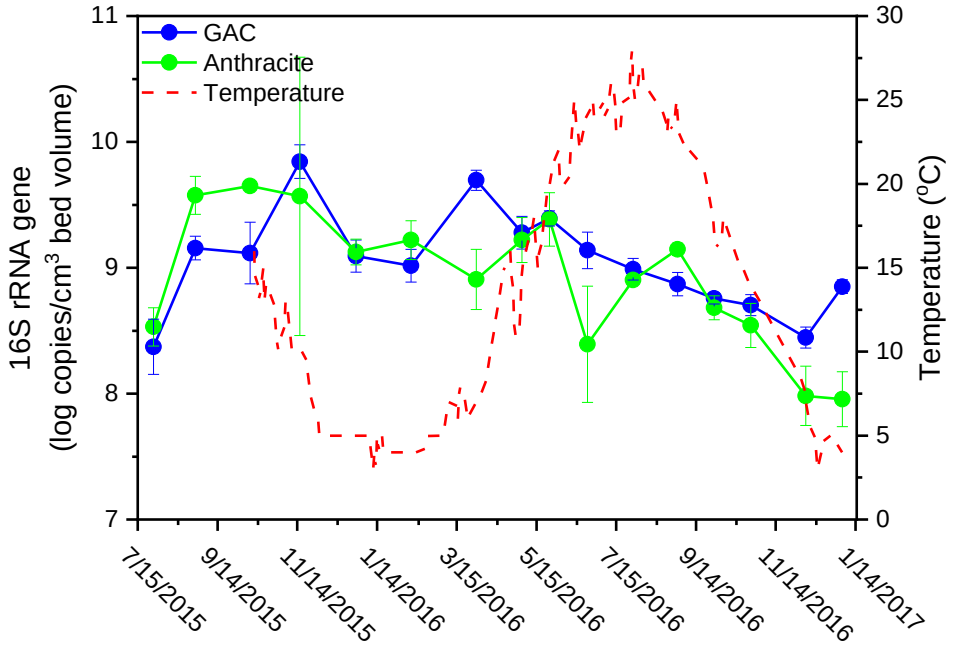

Figure 3- 1. Mean 16S rRNA gene concentration on media samples collected from the top layer of the GAC-sand (blue line; column #1-3) and anthracite-sand biofilters (green line; column #4-6). The red dashed line represents the water temperature. The error bars represent the standard deviation computed from triplicate samples.

3.3.3 Bacterial community composition in biofilters

The bacterial community composition on the media samples and water samples was determined using high-throughput Illumina HiSeq sequencing of PCR-amplified 16S rRNA gene fragments. A total of 124,955,713 high quality sequences were retained, which yielded 56,325 ASVs comprising 1,959 different genera. All ASVs were grouped into

seven categories based on their occurrence in three different types of samples (i.e., media, filter influent and filter effluent) collected from the same biofilters (Table 3-1). The ASVs shared by all three sample types (i.e., media-influent-effluent shared) included the majority of the bacterial populations (>90%). Positive correlations between the mean relative abundance of ASVs and occupancy were observed in all types of samples (Fig. C-4), indicating that "core-satellite" model could be applied in determine the core community. By setting minimal thresholds for occupancy at 50% and mean relative abundance at 0.02%, a total of 632 ASVs were defined as core ASVs, comprising on average 76.1% to 86.1% of the total community in different sample types (Table 3-1). The number of ASVs and their total relative abundance for all detected ASVs and core ASVs are provided in Table 3-1.

The most prominent genera in the media samples were *Hydrogenophaga*, unknown genera in *Rhodobacteraceae* and unknown genera in *Comamonadaceae* (Fig. 3-2: A and B), comprising 8.4%, 6.0% and 5.8% of total bacterial community on average, respectively. Different dominant genera were observed in the water samples (Fig. 3-2: C, D, and E). *Limnohabitans* and uncultured genera of *LD12 freshwater group* were the most prominent genera in the filter influent and effluent samples, comprising an average of 12.8 % and 11.2 % of the total population.

Table 3- 1. Information of ASV classification for all detected ASVs and the core ASVs.

Categories*	Column # 1-3 (GAC-sand biofilter with CECs additions)				Column # 4-6 (Anthracite-sand biofilters with CECs additions)				Column # 7&8 (GAC-sand biofilter without CECs additions)			
All ASVs detected	Number of ASVs	Total relative abundance			Number of ASVs	Total relative abundance			Number of ASVs	Total relative abundance		
		Media	Influent	Effluent		Media	Influent	Effluent		Media	Influent	Effluent
Media Specific	6023	1.0%	0.0%	0.0%	9609	1.3%	0.0%	0.0%	4278	0.9%	0.0%	0.0%
Influent Specific	14040	0.0%	1.1%	0.0%	13398	0.0%	1.0%	0.0%	15445	0.0%	1.3%	0.0%
Effluent Specific	6175	0.0%	0.0%	0.8%	5365	0.0%	0.0%	0.7%	5106	0.0%	0.0%	0.9%
Media-Influent Shared	2102	0.5%	0.3%	0.0%	3462	0.9%	0.6%	0.0%	1952	0.8%	0.5%	0.0%
Media-Effluent Shared	1244	3.9%	0.0%	1.0%	1197	1.7%	0.0%	0.8%	755	4.4%	0.0%	1.4%
Influent-Effluent Shared	2845	0.0%	0.7%	0.8%	1807	0.0%	0.4%	0.5%	2689	0.0%	1.0%	1.2%
Media-Influent-Effluent Shared	4907	94.6%	97.8%	97.4%	5227	96.1%	98.0%	97.9%	3808	93.8%	97.3%	96.6%
Total number of ASVs		14276	23894	15171		19495	23894	13596		10793	23894	12358
Core ASVs	Number of ASVs	Total Relative abundance			Number of ASVs	Total Relative abundance			Number of ASVs	Total Relative abundance		
		Media	Influent	Effluent		Media	Influent	Effluent		Media	Influent	Effluent
Media Specific	225	25.5%			191	18.9%			195	22.4%		
Influent Specific	70		3.9%		61		3.1%		75		4.7%	
Effluent Specific	10			0.4%	9			0.3%	16			0.7%
Media-Influent Shared	23	3.1%	2.5%		42	3.7%	4.4%		16	2.2%	1.3%	
Media-Effluent Shared	39	10.6%		2.5%	18	3.6%		0.7%	43	12.9%		2.9%
Influent-Effluent Shared	61		4.4%	4.2%	35		2.1%	2.3%	72		6.3%	6.1%
Media-Influent-Effluent Shared	139	47.0%	72.5%	71.1%	155	56.7%	73.7%	72.7%	130	48.3%	71.0%	68.5%
Total number of core ASVs	567	426	293	249	511	406	293	217	547	384	293	261
Total relative abundance		86.1%	83.3%	78.2%		82.8%	83.3%	76.1%		85.8%	83.3%	78.3%

***Note:** All ASVs were grouped in to seven categories. The ASVs that were detected only in media, filter influent, or filter effluent samples were defined as ‘media specific’, influent specific’, and ‘effluent specific’. The ASVs that were detected in two different types of samples but not the third type were defined as ‘media-influent shared’, ‘media-effluent shared’ and ‘influent-effluent shared’. The ASVs that were detected in all three types of samples were defined as ‘media-influent-effluent shared’.

The bacterial communities on the GAC and anthracite media samples were highly diverse. Shannon index values ranged from 5.6±0.2 to 8.3±0.0 for the GAC media samples, and 5.6±0.3 to 8.2±0.1 for the anthracite media samples (Table C-2). There was no significant difference (P=0.141) between the Shannon index values for the GAC and anthracite media samples. Also, no clear seasonal change was observed. The diversity of the bacterial community decreased with depth for the GAC and anthracite media samples, as indicated by Shannon index values (Table C-3). The bacterial diversity in the water samples decreased after filtration on most sampling dates, with Shannon index values of 5.9±0.4 to 7.5±0.1 for the filter influent, and 5.8±0.0 to 7.2±0.1 for the filter effluent (Table C-4).

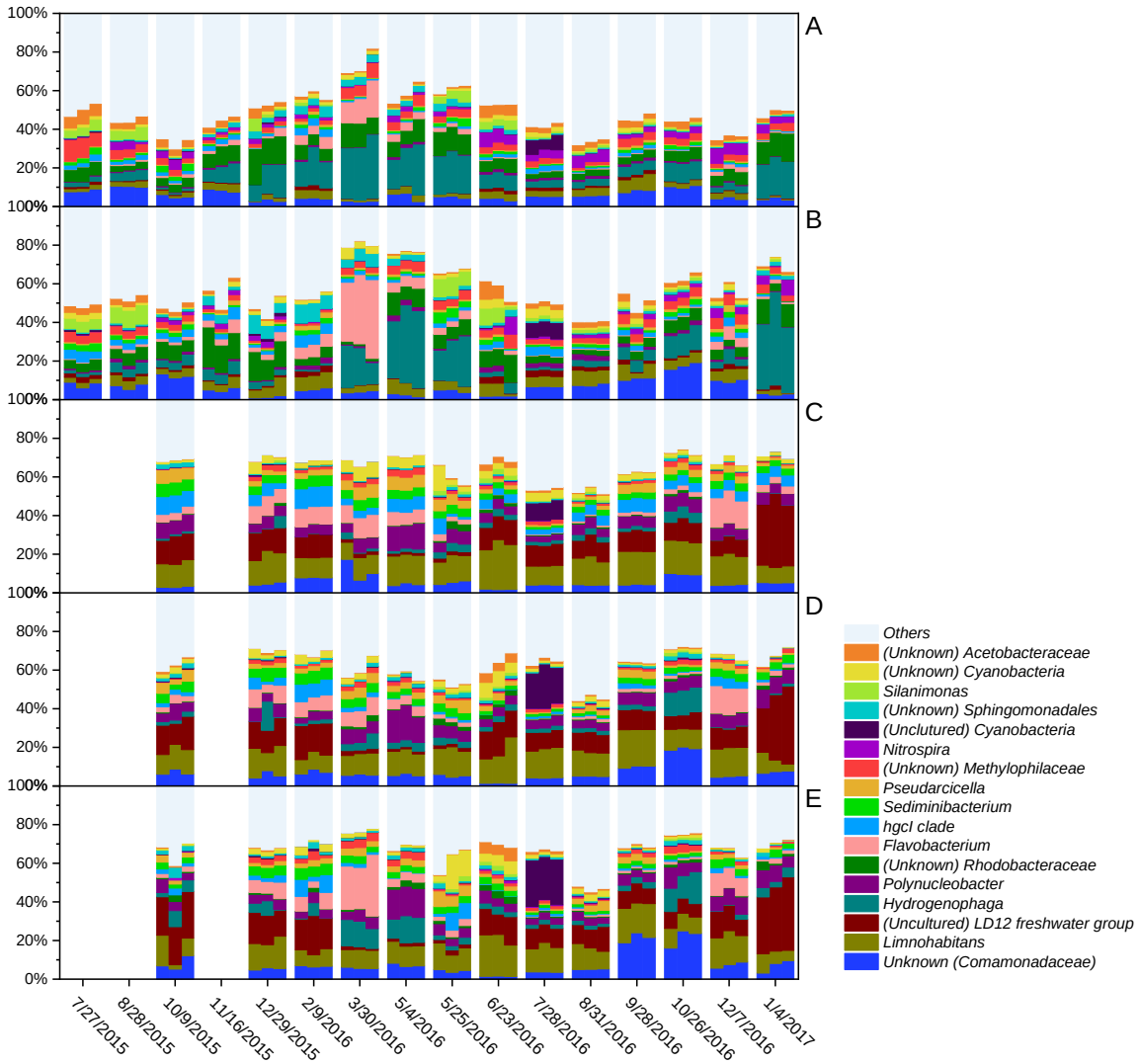

Figure 3- 2. Bacterial community profiles of showing the 17 most abundant genera and unclassified sequences represented in (**A**) GAC media samples (column #1-3) and (**B**) anthracite media samples collected from top layer of biofilters, (**C**) filter influent, (**D**) filter effluent from GAC-sand biofilters (column #1-3) and (**E**) filter effluent from anthracite-sand biofilters. The abundant genera are those that comprised greater than 5% of total sequences in at least 16 of 319 samples in total that were analyzed for DNA sequencing in this study. The genera that are not considered as abundant genera are grouped in “others”. The triplicate samples were plotted separated within the same panel. No water samples were collected on 7/27/2015,8/28/2015 and 11/16/2015

3.3.4 Variations in the bacterial community on the biofilter media

Temporal variation and effect of water quality. PCoA plots revealed distinct clusters based on the sampling dates, suggesting that the bacterial community composition exhibited temporal variations (Fig. 3-3: A, B and C). The first principal coordinates in the PCoA plots are significantly correlated ($P<0.05$, Pearson's correlation) with the water temperature (Fig. C-5), indicating that temperature or water quality parameters that correlate with temperature, such as the concentrations of DOC and ammonia, may explain the temporal variations in the bacterial community. PERMANOVA tests indicated that all three parameters had significant effects on the bacterial community ($P_{adonis}<0.05$; Table C-5). Among these three parameters, temperature explained most of the variations, with R^2_{adonis} of 0.25 and 0.27 for GAC and anthracite media samples, respectively (Table C-5). To determine the impacts of the water quality parameters on individual ASVs, the relative abundance of each ASV and temperature, influent DOC and ammonia concentrations were compared in multiple correlation analyses. Among 567 core ASVs, 63, 26 and 22 ASVs detected in GAC media samples (collected from column #1-3) had statistically significant correlation ($P_{adjusted} < 0.05$) with temperature, influent DOC and ammonia concentrations, respectively (Fig. C-6). Similar results were observed for the anthracite media samples, and GAC media samples without continuous CECs addition (Fig. C-7 and C-8).

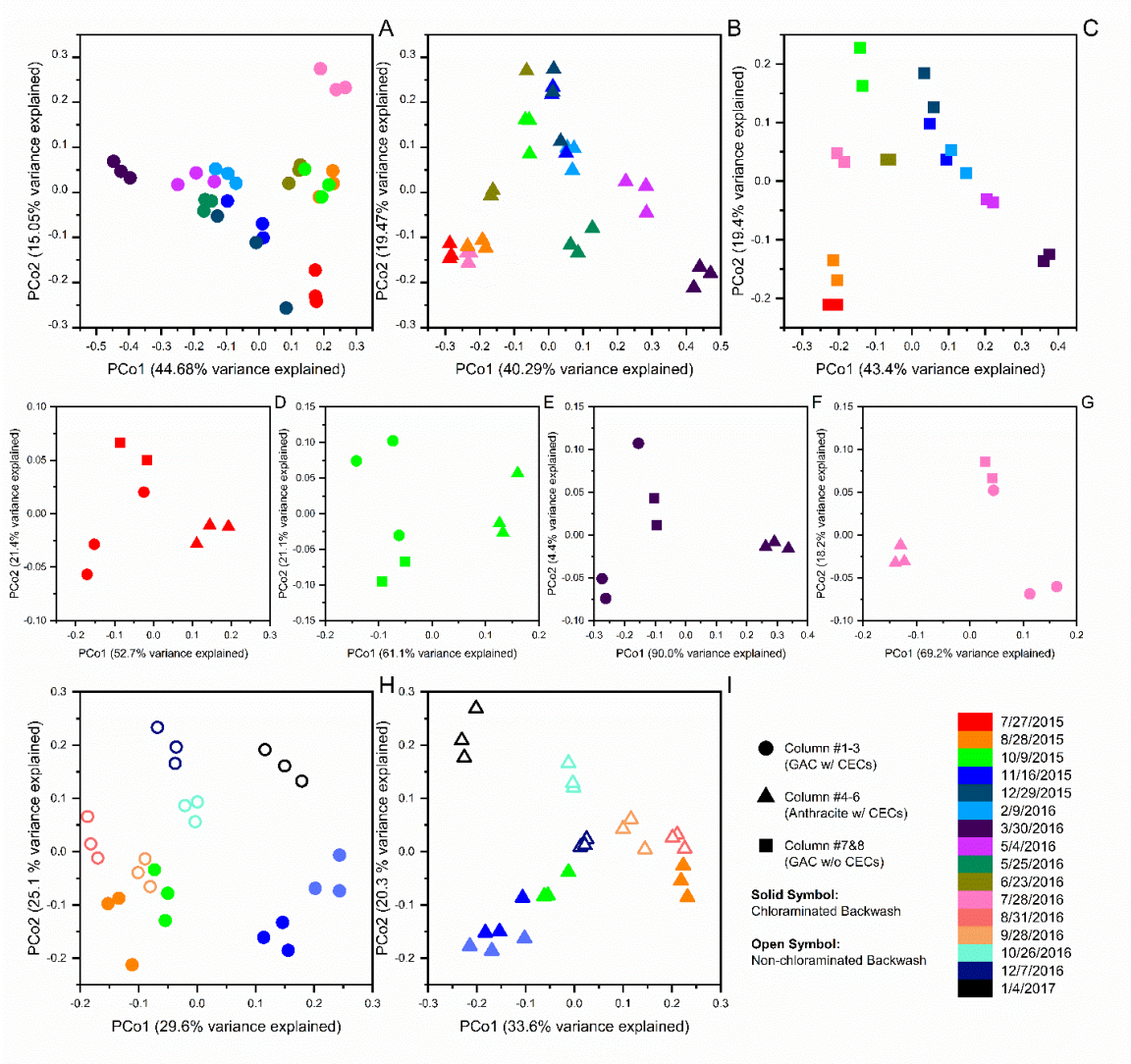

Figure 3- 3. PCoA plots based on weighted UniFrac distance metrics for the media samples collected from the top layer of the biofilters showing: the temporal variation in the bacterial community in the biofilter column #1-3 (**A**, GAC with CECs addition), column #4-6 (**B**, anthracite with CECs addition), and column 7&8 (**C**, GAC without CECs addition); the effect of media types (GAC vs. anthracite) and CECs addition on the bacterial community on the media samples collected on 07/27/2015 (**D**), 10/09/2015 (**E**), 03/30/2016 (**F**), and 07/28/2016 (**G**). The results from the rest sampling dates are shown in Fig. C-9; the effect of backwash strategies (chloraminated vs. non-chloraminated) on the bacterial community in the biofilter column #1-3 (**H**) and column #4-6 (**I**). Different color represents the sampling date.

Effect of media type. Distinct clusters of GAC and anthracite samples in the PCoA plots were constantly observed on each sampling date (Fig. 3-3: D, E, F, and G, and Fig C-9), suggesting that there was a significant effect of media type on the bacterial community composition in the biofilters. PERMANOVA testing indicated that the bacterial communities in the GAC and anthracite biofilters were significantly different (R^2_{adonis} =0.09, P_{adnois} =0.001; Table C-6). None of the bacterial communities in the triplicate GAC and anthracite media samples collected on each sampling date, however, were significantly different according to pairwise PERMANOVA tests ($P_{adnois\_adjusted}$ = 0.12 to 0.31; Table C-6). Differential abundance analyses indicated that the relative abundances of 170 core ASVs in the GAC and anthracite media samples were significantly different ($P_{walt\_adjusted}$<0.05; Fig. C-10).

Effect of continuous CECs addition. No significant difference in the bacterial communities was observed in the GAC-sand biofilters with and without continuous trace amount of CECs addition, as supported by PCoA plots (Fig. 3-3: D, E, F, and G, and Fig C-9) and PERMANOVA tests (p_{adnois}=0.88; Table C-7). The differential abundance analyses indicate that only 23 core ASVs differed significantly in relative abundance on the GAC media samples with and without CECs addition (Fig. C-11).

Effect of backwash. The bacterial communities on the media samples collected during the same season, but one year apart, under different backwashing conditions (08/15-01/16: chloraminated backwash; 08/16-01/17: non-chloraminated backwash) were compared in

PCoA plots (Fig. 3-3: H and I). The bacterial community exhibited a significant shift after switching the backwash strategy, with P_{adnois} of 0.001 and 0.002 for GAC and anthracite media samples, respectively. Differential abundance analyses indicated that the relative abundance of 285 core ASVs were significantly different ($P_{walt\_adjusted}$<0.05) on the media samples with chloraminated and non-chloraminated backwash (Fig. C-12).

Depth variation. The bacterial community exhibited variations with depth in the biofilters (Fig. 3-4). Pairwise PERMANOVA tests indicate that the bacterial communities in the top layer (0 m) were significantly different ($p_{adonis\_adjusted}$ <0.05) from those at greater depths in the GAC-sand and anthracite-sand biofilters (Table C-8).

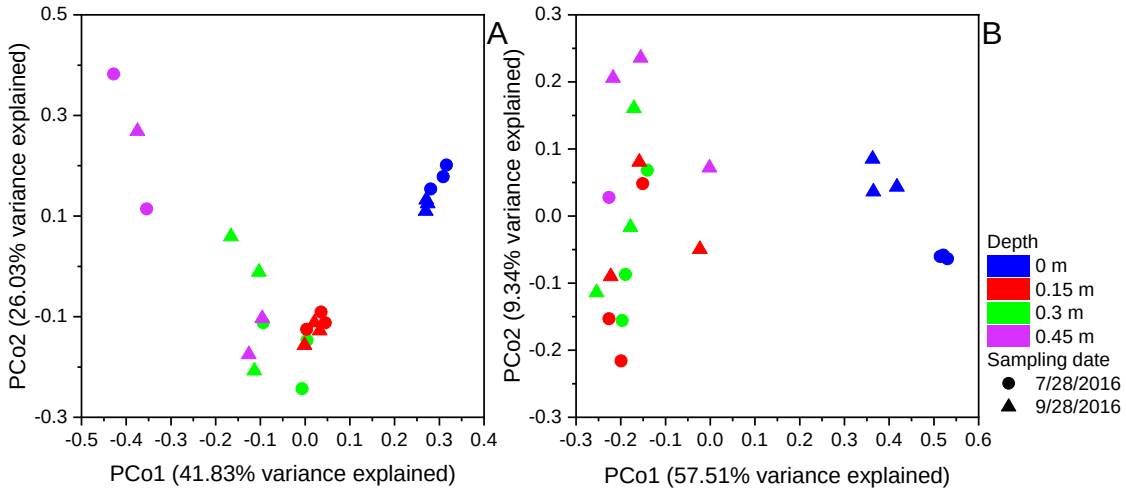

Figure 3- 4. PCoA plot based on weighted UniFrac distance metrics showing the variations in the bacterial community on GAC (**A**, column #1-3) and anthracite (**B**, column #4-6) media samples collected on 07/28/2016 and 09/28/2016 from four different depths (0 m, 0.15 m, 0.30 m, and 0.45 m) in the biofilters.

3.3.5 Effect of biofilter on the bacterial communities in filter effluent.

PCoA plots were prepared to visualize the difference in bacterial communities between filter influent, filter media (top layer) and filter effluent samples collected from the GAC-sand and anthracite-sand biofilters (Fig. 3-5: A and B). There was clear separation between water samples and media samples in the PCoA plots, suggesting that the bacterial communities in the water samples and the media samples were different. Furthermore, the close proximity of the filter influent and effluent samples suggested that those communities were similar. Such observations were also supported by statistical tests (pairwise PERMANOVA, Table C-9). The bacterial communities in the media samples were significantly different ($p_{adonis\_adjusted}$ <0.05) from those in the filter influent and effluent samples, whereas no significant difference ($p_{adonis\_adjusted}$ >0.05) was observed between the bacterial communities in the filter influent and effluent.

To further assess intra-ASV level difference in the bacterial communities, the mean relative abundance values of all detected ASVs were compared between biofilter media, filter influent, and effluent (Fig 3-5: C and D). The greatest Pearson's correlation coefficients were obtained for comparisons of the filter influent and effluent (0.97 and 0.97 for GAC-sand and anthracite-sand biofilters, respectively), supporting the previous statement that the communities in those samples were similar. The correlation coefficients between the filter effluent and media were less than those between the water samples, but constantly greater than between the filter influent and media, suggesting there was a clear, albeit

minor, impact of the biofilters on the bacterial communities in filter effluent. Similar results were also observed by comparing the community dissimilarities using the weighted UniFrac distance metrices (Fig. C-13). Relatively lower dissimilarity values were observed between the filter effluent and media than between the filter influent and media. The mean relative abundance of all core ASVs in the filter media, influent and effluent samples collected from GAC-sand (column #1-3) and anthracite-sand (column #4-6) biofilters are also provided (Fig. C-14 and C-15).

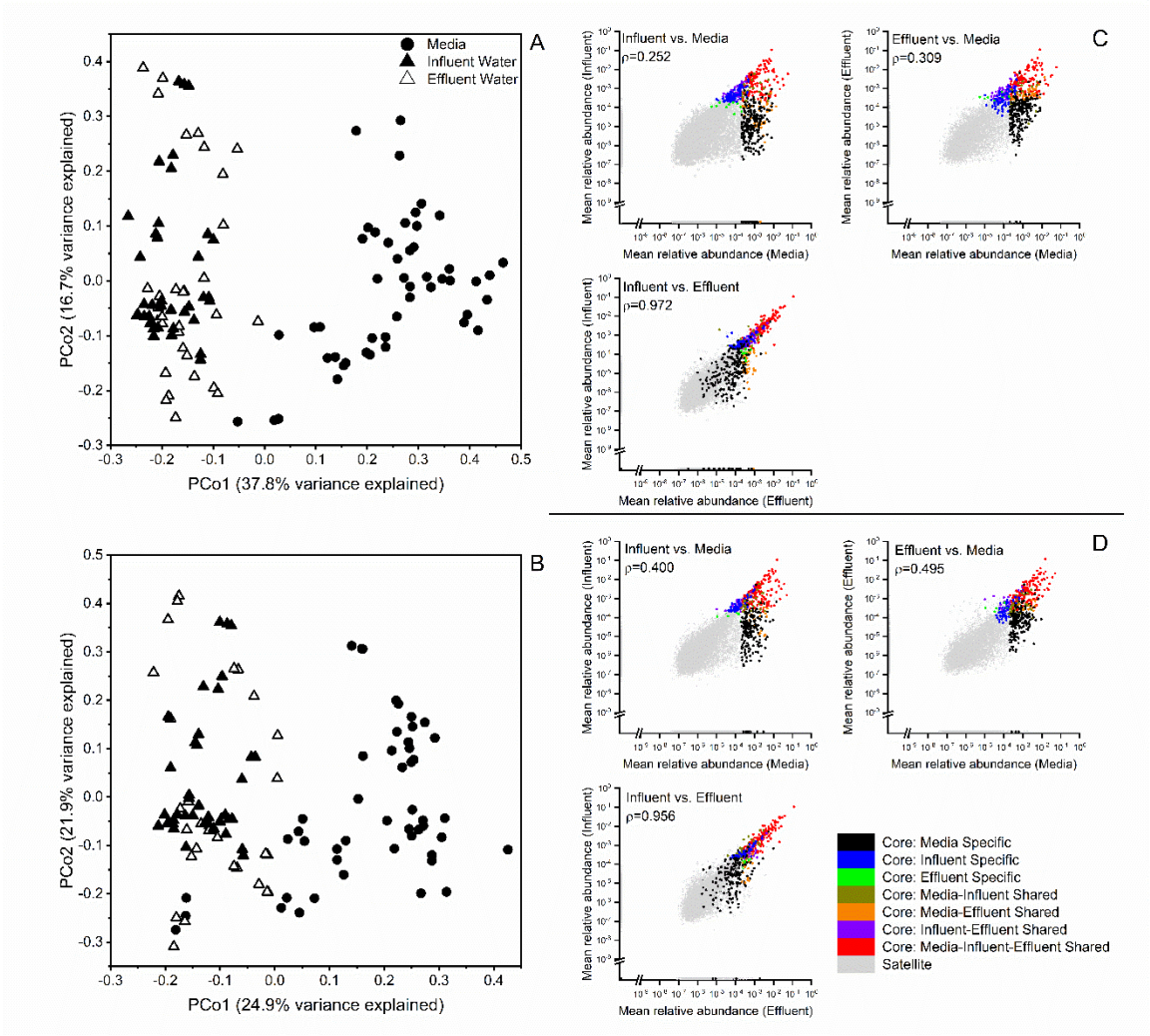

Figure 3- 5. Comparison of the bacterial communities in biofilter media, filter influent, and effluent samples. PCoA plots based on weighted UniFrac distance metrics showing the bacterial community composition in the samples collected from the top layer of the GAC-sand biofilters (**A**, column #1-3) and anthracite-sand biofilters (**B**, column #4-6); Mean relative abundance of all detected ASVs in the samples collected from the top layer of the GAC-sand biofilters (**C**, column #1-3) and anthracite-sand biofilters (**D**, column #4-6). ρ is the Pearson's correlation coefficient value comparing the mean ASVs relative abundance between two types of samples.

3.4 Discussion

This study provided a comprehensive investigation of the bacterial community in the pilot-scale biofilters. A work flow, including community visualization (bar charts and PCoA plots. Fig. 3-2 to 3-5), core population identification and classification (Table 3-1), and multiple statistical tests (PERMANOVA, multiple correlation analysis, and differential abundance analysis) were used to explore temporal and spatial variations, the effect of media type, CECs addition and backwash strategy on the biofilter microbiome, and the interaction of the microbiome on the media with that in the water. This study should be beneficial in understanding the microbiome in water treatment biofilters from both ecological and engineering perspectives.

A unique aspect of the experimental setup in this study was the operation of triplicate (column #1-3 and 4-6) and duplicate (column 7&8) biofilters columns in parallel for an extended period of time (18 months). High reproducibilities of sequencing results among the replicate samples were consistently observed (Fig. 3-2 to 3-4), which ensured the statistical assessments of variations and differences in the bacterial community under different conditions (e.g., media type and CECs addition). Relatively higher dispersions among the replicate media samples collected at greater depth (Fig. 3-4) were expected as the environmental conditions, such as DOC and dissolve oxygen (DO), changed along the biofilters. Such changes may deviate among replicate biofilter columns.

The bacterial community in biofilters shared similarities with other ecosystems. The prominent genera, such as *Hydrogenophaga*, *Limnohabitans*, and *Flavobacterium*, are commonly found in freshwater environments [72, 73], soil [74], and water distribution systems [14, 68, 69, 75]. Human gut bacteria (e.g., *Bacteroides* [76]) and bacteria from genera that include opportunistic pathogens (e.g., *Legionella* and *Mycobacterium*; [77]) were also detected in the biofilters as minor populations, comprising less than 0.2% of the total community. As a part of the source-to-tap microbiome, the present of these bacteria in the biofilters may pose potential risks to drinking water safety and public health.

The total biomass level did not change significantly over time despite seasonal variations in water temperature and influent DOC concentration (Fig. 3-1). This is unexpected because higher temperature and influent DOC concentration should be beneficial for growth of the mesophilic heterotrophic bacteria that tend to dominate the communities of water treatment filters [78]. The overall biomass in a biofilter at a given point in time is a function of the rates of growth, deposition, and loss by physical removal via fluid shear as well as decay, inactivation, and predation, with increasing biomass expected during periods where the combined growth and deposition rates exceed the loss rate [79, 80]. It is possible that the combined rates of growth and deposition remain in balance with the loss rate throughout the year. Although the total biomass level was stable, the bacterial community composition exhibited temporal variation that was likely due to water quality changes. Bacteria respond to such changes differently according to their physiological characteristics. For example, the relative abundance of psychrophilic bacteria

Polaromonas was negatively correlated with temperature in all biofilters, whereas the relative abundance of thermophilic bacteria, such as *Limnohabitans* and *Silanimonas*, tended to increase in warmer temperatures (Fig. C-6 to C-9) [81–83]. Considering all three parameters discussed in the present study (temperature, influent DOC and ammonia concentrations), PERMANOVA results suggested that more than 50% variance in the bacterial community was not explained by these parameters (Table B-5). Thus, other water quality factors, such as DO, phosphate, sulfate, and natural organic matter composition may have also affected the community composition [14]. The community composition on the media samples collected at the same time of year (07/27/2015 vs. 07/28/2016) were relatively similar as indicated in the PCoA plots (Fig. 3-3: A, B and C), suggesting that the temporal variations could be seasonal. Similar water quality driven temporal variations in the bacterial community were observed in biofilters and distribution systems in previous investigations [13, 68, 69].

Long-term (>1 year), continuous CECs addition did not appear to significantly affect the bacterial community in the biofilters, which disagreed with most previous studies investigating the effect of CECs on bacterial communities in various natural and engineered ecosystems [58, 59, 84]. Such disagreement is likely attributed to the difference in the applied CECs concentrations. Environmentally relevant trace levels of CECs (0.1-3 µg/L) were spiked into the biofilter influents in the present study, whereas much higher concentrations (> 10 µg/L) were applied in most previous studies. The EC_{50} values (i.e., the effective concentration to reach half of the maximum response in bacteria) of all

selected CECs ranged from 54 to 35,440 µg/L [85–88]. Minor differences in the relative abundances of few ASVs were observed on the GAC media with and without CECs addition (Fig. C-12). There is, however, no direct connection between those ASVs and CECs resistance or survival mechanisms.

The effect of backwash strategy on the bacterial community in biofilters provides novel insights on how the bacterial community responds to periodic, short-term contact of disinfectant (i.e., chloramine). Our results suggest that disinfectant can still significantly impact the bacterial community, even with 10-minute application every 4 to 6 days. During backwash, the biofilms on the filter media were also disturbed by the high velocity water [89, 90], which may amplify the effect of disinfectant. Interestingly, mycobacteria and AOB, which were previously reported as bacteria well-adapted to chloramine [60, 61], did not appear to increase in abundance under chloraminated backwash condition (Fig. C-13). It is important to note that the comparison between two backwash strategies were performed in the same season but one year apart. It can be argued that the differences in bacterial community under different backwash strategies were the result of year-to-year variations, even though the temporal variation was likely to be seasonal, as discussed above. In parallel comparison is recommended to address such argument.

The difference in bacterial communities between GAC and anthracite media sample was likely due to the distinct surface characteristics of the two media (e.g., porous structure and hydrophobicity), which may affect the colonization of microorganisms and carbon

availability [15, 89]. Also, GAC can catalytically reduce chloramine [91], which was applied in the backwash water for the first year of the operation. The difference in the catalytical properties between the two media may also result in the community difference, especially considering there was a significant effect of chloramine on the bacterial community in biofilters observed in the present study (i.e., the effect of backwash strategy).

The reductions of total biomass and diversity of the bacterial community with depth were expected because there were less resources available (i.e., assimilable organic carbon, DO, and nutrients) for bacterial growth. Similar results were observed in biofilters [15, 92] as well as other ecosystems (e.g., soil and sediment [93, 94]). The bacterial community also varied with depth, probably due to water quality changes and the effects of bacterial deposition from the water onto the media, which was expected to be more important near the top of the filter bed. The bacterial community in the top layer is critical in determining the biological functions of the biofilters as it contains the most biomass, and the high diversity ensures the performance stability and resistance to process upsets [95, 96].

The suspended community (water) and biofilm community (media) were significantly different despite that they were within the same ecosystem. Similar results were also observed in other ecosystems, such as water distribution systems and watershed [68, 97, 98]. Such difference was probably attributed to the distinct biofilm formation related characteristics between microorganisms, such as extracellular polymeric substances (EPS) characteristics and quorum sensing [99–101]. Interestingly, one previous study showed that

Hydrogenophaga was dominant in the water but not the biofilm in the drinking water distribution system [68], whereas it was found as one of the prominent population in the media samples (i.e., biofilm community) in the present study. This suggests that the difference between the suspended community and biofilm community may vary in different ecosystems, probably due to distinct intraspecies competition as the bacterial communities are different.

The impact of biofilters on the microbiome in filter effluent was consistent with previous reports in the literature [14, 15]. Biofilters impact the water community through collecting biomass from the filter influent and shedding bacteria into the filter effluent. The former process appeared to be more important in our system, as exhibited by approximately 70% reduction in biomass concentration in the water from influent to effluent. Interestingly, there was very little change in community composition, suggesting that the probability of being filtered out was similar for all populations and that bacteria shed from the filter were a minor component of the effluent communities. Grouping bacteria based on their abundance and occupancy in different samples highlights the organism specific selection in collecting and shedding. 'Influent specific' bacteria like *Limnohabitans* were more likely to be captured by the biofilters than bacteria in other groups, whereas shedding from the biofilters was the primary source of 'media specific' and 'media-effluent shared' populations, such as *Bdellovibrio* and nitrifying bacteria (*Nitrosomonas spp.* and *Nitrospira spp.*), in the filter effluent (Fig. C-14 and C-15). Such selection was likely attributed to the removal mechanisms and biofilter stability (e.g., shear resistance) on the

media surface, which are related to the bacterial physiological characteristics such as the cell size, EPS characteristics and cell surface hydrophobicity [99, 102].

In summary, this study provides a comprehensive understanding of microbiome in biofilters. The bacterial community in the biofilters exhibited temporal and spatial variations due to the water quality changes over time and varied along the filter bed. The filter media type and the backwash strategy significantly impacted the bacterial community, which provides potential methods of managing the biofilters microbiome, possibly for better contaminant removal performance and opportunistic pathogen control. The biofilters can effectively remove biomass from the water, but only marginally impact the post-filtration bacterial community composition in the water, suggesting that the ecological impact of the biofilters should also be considered in the future biofilter design and operation. The work flow in the present study should be useful in analyzing the microbiome in other complex ecosystem, especially those containing both suspended and biofilm community, like riverbank and drinking water distribution system. Future research should explore the impact of other water quality parameters, such as DO, phosphate and nitrate, and upstream treatment processes on the biofilter bacterial community. Detailed functional analysis of metagenomes and metatranscriptomes is also recommended to further explore the biological functions of the biofilters.

3.5 Acknowledgement

Financial support for this work was provided by the City of Minneapolis. This work was carried out in part using Illumina HiSeq instruments at Genomic Center at University of Minnesota, Twin Cities. We thank Annika Bankston, George Kraynick and Dr. Li Zhang of Minneapolis Water Treatment and Distribution Services for helping coordinate this study and providing technical input. We thank Lucian Osuji, Christopher Rydell, Andy Weyer, Joe Kroening, Troy Rosenthal and Eric Raway for helping with sample collection and water quality analysis. We also thank Dr. Michael Waak for his inputs in the DNA sequencing analytical methods.

Chapter 4

Investigation of the microbiome in fourteen full-scale biofilters across North America

A key objective in ecology is to understand how biodiversity varies across space. Spatial patterns of bacterial communities often follow a distance-decay relationship in which community similarity decreases with geographic distance. Such a relationship has been commonly observed in a variety of natural ecosystems, such as freshwater, soil, and salt marshes, but not in engineered ecosystems. In this study, biogeographic patterns of the microbiomes in full-scale biofilters across North America were investigated. Bacterial abundance and community composition on the filter media samples (n=57) from biofilters at fourteen water treatment facilities across North America were determined by analyzing 16S rRNA genes using real-time quantitative polymerase chain reaction and Illumina HiSeq high-throughput sequencing. Bacteria were abundant on the filter media (8.75 ± 0.29 to 10.67 ± 0.18 log 16S rRNA gene copies/cm^3 bed volume) and the bacterial communities were highly diverse (Shannon index: 5.30±0.06 to 8.43±0.01). Significant inter-filter variations were observed that followed a weak, but highly significant, distance-decay relationship ($z=0.00573\pm0.00058$; $P=1.78\times10^{-22}$). Bacterial community composition correlated more strongly with the water quality characteristics at the time of media sampling (e.g., pH, temperature, influent dissolved organic carbon concentration) than with

geographic distance according to a multiple regression on matrix analysis. A Mantel test, however, indicated that geographic distance was significantly correlated with all three water quality parameters ($P_{mantel} < 0.05$). This study has provided some novel insights into the factors that shape the microbiomes of biofilters.

4.1 Introduction

The effects of environmental variables and geographic distance on bacterial community composition in ecosystems in many natural environments, such as freshwater, soil, and salt marshes, have been well documented [93, 103–105]. Typically, bacterial community similarity decreases with geographic distance, which has come to be known as the distance-decay relationship [106]. Such a relationship could be explained by differences in environmental variables, such as temperature, because these variables are often autocorrelated with geographic distance. Hence, this observed environmentally-driven distance-decay relationship for bacterial communities agrees with the Baas-Becking hypothesis wherein "everything is everywhere, but the environment selects".

Little is known about the biogeographic patterns of bacterial communities in engineered ecosystems. Previous studies have suggested that the microbiomes in wastewater treatment bioreactors followed a similar but weaker distance-decay relationship compared to natural ecosystems [107, 108]. Biofilters, which are granular media filters (i.e., sand, anthracite coal, and/or granular activated carbon) inhabited by viable bacteria, have been widely used for drinking water treatment since the 1800s [12]. There are more than 14,000 drinking water treatment systems in the United States using surface water as their source water, with most using biofiltration in their water treatment processes whether intentional or not (Safe Drinking Water Act, 1986). The microbiome in biofilters can be beneficial for degrading various contaminants, including natural organic matter [8, 9], taste-and-odor compounds

[10], and other micropollutants [11]. In addition, it is important to consider biofilters as part of the 'source-to-tap' microbiome, because biofilters influence the microbial communities in the finished water leaving the treatment facility [14]. Understanding bacterial biogeographic patterns in biofilters is important as it considers a new twist on the ecological question of nurture versus nature. Namely, which factors have a stronger influence on the microbial communities that develop in engineered biofilters, 'nature' factors such as the source water microbiome and water quality or engineering decisions (i.e., 'nurture' factors) such as the choice of filter media and prefiltration treatment operations. Such information could prove valuable for enhancing or maintaining biofilter performance as well as limiting potential negative effects of biofilters on treated water quality.

In this study, the bacterial abundance and community composition of full-scale biofilters at fourteen drinking water treatment facilities across North America were investigated by quantifying 16S rRNA genes using real-time quantitative polymerase chain reaction (qPCR) and sequencing the V3 region of the 16S rRNA gene on the Illumina HiSeq platform. The effects of geographic location (i.e., inter-facility distance), contemporaneous water quality variables (e.g., pH, temperature, and dissolved organic carbon (DOC) concentration), and engineering decisions (e.g., media type and prefiltration ozonation application) on full-scale biofilter bacterial communities were evaluated. Finally, a multiple regression on matrix (MRM) analysis was performance to identify the relative importance of each abovementioned factor.

4.2 Method

4.2.1 Sample collection and DNA extraction

Filter media samples were collected from full-scale filters at fourteen different water treatment facilities across North America. The biofilters were selected to provide a range of influent water qualities, media types (GAC or anthracite), and pre-filtration treatment processes (e.g., with and without ozonation; Table 4-1). All biofilters were initially sampled between March 2017 to June 2017 with five facilities (C, E, G, I, and J) sampled for a second time in August 2017. Influent water quality parameters, including pH, water temperature and influent DOC concentration, were also measured on the day of sample collection (Table 4-1). In total, 57 biofilter media samples (i.e., triplicate samples from 19 sampling events) were collected.

Triplicate filter media samples (approximately 0.5 g as wet weight each) were collected in sterile microcentrifuge tubes from the top 7.5 cm of each filter bed. All media samples were frozen and then shipped on ice to the laboratory via overnight carrier. The media samples were subjected to three consecutive freeze-thaw cycles to lyse the cells and then the genomic DNA was extracted using the FastDNA SPIN Kit for Soil (MP Biomedicals, Solon, OH). Extracted DNA was stored at -20 °C until future use. After the DNA was extracted, the media in each microcentrifuge tube was removed, dried at 105 ^{o}C, and then weighed to obtain the dry weight of the media. The corresponding media volumes were

Table 4- 1. Information of sample collection, participating facilities and biofilter influent water qualities (measured on the day of sample collection).

Sample ID	Biofilter ID	Geographic Location	Sampling Date	Media	Ozone	pH	Temperature (°C)	Influent DOC concentration (mg C/L)	Source Water	Coagulant
A	A	CA	3/15/2017	GAC	Yes	7.07	12.3	N/A[1]	Reservoir	Aluminum[2]
B	B	ON	5/15/2017	GAC	Yes	7.73	13.8	4.03	River	Aluminum
C1	C	TX	5/17/2017	GAC	Yes	7.4	23.1	3.44	Lake	Aluminum
C2	C	TX	8/8/2017	GAC	Yes	7.48	28.9	3.67	Lake	Aluminum
D	D	TX	8/7/2017	GAC	Yes	7.46	30.9	3.72	Lake	Aluminum
E1	E	MI	6/5/2017	GAC	Yes	9.3	21.1	4.26	River with option for groundwater	None
E2	E	MI	8/22/2017	GAC	Yes	9.5	24.3	N/A	River with option for groundwater	None
F	F	GA	6/12/2017	Anthracite	Yes	6.83	19.3	1.2	Lake	Ferric[3]
G1	G	MN	5/1/2017	GAC	No	8.7	11.2	4.65	River with option for groundwater	Aluminum and Ferric
G2	G	MN	8/11/2017	GAC	No	8.86	24	4.01	River with option for groundwater	Aluminum and Ferric
H	H	VA	5/2/2017	Anthracite	Yes	6.5	24	2.6	Reservoir	Aluminum
I1	I	CO	6/6/2017	GAC	No	7.26	15.1	2.56	Blended	Ferric
I2	I	CO	8/2/2017	GAC	No	7.5	17.9	2.2	Blended	Ferric
J1	J	CO	6/6/2017	GAC	No	6.63	11.3	2.18	Reservoir	Ferric
J2	J	CO	8/2/2017	GAC	No	6.48	14	2.09	Reservoir	Ferric
K	K	NC	6/13/2017	GAC	Yes	8.09	25	3.45	River	Aluminum
L	L	NY	6/5/2017	GAC	No	6.1	14.2	1.98	Reservoir	Aluminum
M	M	TX	6/13/2017	Anthracite	Yes	6.9	N/A	3.28	Lake	Ferric
N	N	TX	6/13/2017	GAC	Yes	7.5	N/A	3.24	Lake	Ferric

Note: 1. Not available; 2. Aluminum based coagulants include alum, polyaluminum chloride and aluminum sulfate;
3. Ferric based coagulants include ferric chloride and ferric sulfate; 4. Sodium hydroxide

computed from the dry weights using typical media bulk density values and then the volumes were used for normalization of the qPCR results.

4.2.2 Real-time quantitative polymerase chain reaction (qPCR)

The 16S rRNA gene for *Bacteria* was used to quantify total bacterial biomass using real-time quantitative qPCR (primers 341F and 518R) according to the protocols of LaPara et al. (2015)[13]. Standards for qPCR were prepared from serial dilution of synthetic double-stranded DNA fragments (gBlocks) to develop calibration curves for quantification of gene copies in each sample. The censoring limit (CL) was defined as the lowest standard of a given qPCR assay that amplified (1100 copies per reaction), normalized by the volume (L) of the water sample or the bed volume (cm^3) of the media sample used for DNA extraction. The detailed information of the gene targets and respective primer sequences and gBlock standards is shown in Table C-1. Pearson's correlation test was conducted to determine if a statistically significant correlation existed between two results (e.g., gene concentration and water temperature).

4.2.3 High-throughput 16S rRNA gene sequencing

The DNA extract from each sample was amplified using PCR, purified, quantified, and then pooled by equal mass to create a library for sequencing, as previously described [13]. Next generation high-throughput sequencing of PCR-amplified 16S rRNA gene fragments

(V3 region) was performed on the Illumina HiSeq platform at the University of Minnesota Genomics Center (UMGC) using primers 341F and 534R [63]. Unprocessed sequence reads have been uploaded to Sequence Read Archive (accession PRJNA521294).

The sequence reads were demultiplexed, trimmed, and filtered using QIIME2 (version 2018.2). Amplicon sequence variants (ASVs) were determined using 'DADA2' [64, 65] and then assigned consensus taxonomy using SILVA rRNA database (release 128; [66]). A phylogenetic tree was generated using 'FastTree2' and the rooted using midpoint [67].

Both alpha (Shannon index) and beta (weighted UniFrac distance metrics) diversity were calculated using QIIME2 (pipeline: core-metrics-phylpenetic) after sequence libraries were randomly trimmed down to 100,000 sequences per profile. Principal coordinates analyses (PCoA) plots were made to visualize the weighted UniFrac distance metrics. The distance-decay relationship was determined by comparing the weighted UniFrac similarities (*S*) and the differences in geographic distance (*D*) between the biofilters. The power-law exponent z value was estimated based on the log-transformed equation $\log(S) = \text{constant} - 2z\log(D)$ using *lm* function in R, as previously described [104, 109]. Permutational multivariate analysis of variance (PERMANOVA) tests (999 permutations) were conducted using the *adonis* function in R software to determine the effects of contemporary environmental variables and engineering designs on the bacterial community. The dispersion of the bacterial community within a sample group was tested using the *betadisper* and *permutest* functions in 'vegan' (999 permutations).

4.2.4 Multiple regression on matrix analysis

A multiple regression on matrix (MRM) analysis was performed to identify the relative importance of geographic location, media type, prefiltration ozonation, and water quality parameters (pH, temperature and influent DOC concentration) on bacterial community composition in the full-scale biofilters. The coefficients (B_i) of the following linear MRM relationship were solved for using the *ecodist* package in R [105, 110]:

$$\boldsymbol{UniFrac} = B_0 + B_1\,\boldsymbol{D_{std}} + B_2\,\boldsymbol{pH} + B_3\boldsymbol{T_m} + B_4\boldsymbol{DOC_{inf}} + B_5\,\boldsymbol{M} + B_6\,\boldsymbol{O} \quad (4-1)$$

where: B_i are the regression coefficients; *UniFrac* is the weight UniFrac distance matrix representing the bacterial community dissimilarity between samples; D_{std}, *pH*, T_m and DOC_{inf} are the standardized Euclidean distance matrices representing the normalized differences in the geographic distance between facilities, pH, temperature, and influent DOC concentration on the day of media sample collection, respectively; *M* and *O* are the distance matrices representing the differences in media type and ozonation between the biofilters, where a value of 1 indicates a different media type or ozonation condition and 0 when it is the same. The p-value for each regression coefficient was also calculated. Parameters with no significant effect on the weighted UniFrac dissimilarity at the 95% confidence level were excluded from the final regression equation. Mantel's correlation test (*mantel* function in 'ecodist' package in R; two-tailed) was also performed between the D_{std} and *pH*, T_m and DOC_{inf} to assess if there was a significant correlation between geographic distance and each water quality parameter.

4.3 Results

4.3.1 Quantification of total bacterial biomass

The total bacterial biomass on the full-scale biofilter media ranged from 8.75±0.29 to 10.67±0.18 log 16S rRNA gene copies/cm^3 bed volume (Table 4-2). The 16S rRNA gene concentration correlated positively with pH and influent DOC concentration, with Pearson's correlation coefficients of 0.59 (P=0.01) and 0.74 (P=0.00), respectively. No significant correlation was observed between the 16S rRNA gene concentration and water temperature (P=0.59). The total biomass in the biofilters was relatively stable over time as less than 0.5 log copies/cm^3 bed volume difference in the 16S rRNA gene concentration was observed on the media samples collected three months apart from five selected biofilters (C, E, G, I, and J).

4.3.2 Bacterial community composition in the full-scale biofilters

The bacterial community composition on the filter media was determined using high-throughput Illumina HiSeq sequencing of PCR-amplified 16S rRNA gene fragments. A total of 19,051,124 high quality sequences were obtained, which yielded 12,588 ASVs comprising 984 genera.

Table 4- 2. 16S rRNA gene concentration and Shannon index (Mean ± S. D.) for bacterial communities on the full-scale biofilter media samples.

Sample ID	16S rRNA gene concentration (log copies/cm^3 bed volume)	Shannon Index
A	10.05±0.17	6.50±0.33
B	10.67±0.18	6.17±0.06
C1	10.49±0.04	7.37±0.08
C2	10.05±0.12	7.50±0.05
D	10.02±0.06	7.16±0.04
E1	10.09±0.02	6.56±0.04
E2	10.02±0.06	6.58±0.10
F	9.42±0.77	7.65±0.31
G1	10.16±0.03	8.03±0.01
G2	10.16±0.11	8.43±0.01
H	8.75±0.29	7.66±0.00
I1	9.40±0.02	5.73±0.08
I2	9.68±0.05	5.30±0.06
J1	9.21±0.03	6.67±0.04
J2	9.26±0.04	6.48±0.28
K	9.45±0.06	5.90±0.12
L	9.16±0.09	5.95±0.06
M	9.77±0.36	7.84±0.21
N	10.02±0.06	7.60±0.08

The bacterial communities in the full-scale biofilters were highly diverse. Shannon index values ranged from 5.30±0.06 (Biofilter I) to 8.43±0.01 (Biofilter G) (Table 4-2). No significant correlation was observed between Shannon index value and pH, temperature, or influent DOC concentration (P=0.60, 0.62, and 0.19, respectively). The bacterial community diversity was relatively stable over time as the differences in Shannon index values for the media samples collected from the same biofilters on the two sampling dates were less than 0.5.

The most prominent genera in the biofilters were: unknown genera in *Comamonadaceae*, unknown genera in *Sphingomonadales*, *Bradyrhizobium*, and unknown genera in *Blastocatellaceae* (Fig. 4-1). These populations comprised on average 6.6%, 5.4%, 3.7% and 3.6% of the total bacterial community in all biofilter media samples, respectively, and these populations were detected in more than 98% of the media samples (Table D-1). Other prominent genera, such as *Nitrospira* and *Hydrogenophaga*, were also frequently detected but comprised on average less than 3% of the total bacterial community in the biofilters (Table D-1).

PCoA plots revealed distinct clusters for each biofilter, suggesting that there were significant inter-filter variations in bacterial community composition (Fig. 4-2). Temporal variations in bacterial community composition were also observed by comparing media samples collected from the same biofilters three months apart but were relatively smaller than the inter-filter variations. Clustering in the PCoA plots also seemed to be affected by the geographic locations of the biofilters. For example, the community compositions of two biofilters located in Texas (M and N; Table 4-1) were relatively similar as indicated by the close proximity of the data points in the PCoA plots. Conversely, those data points were distant from points representing the communities in a biofilter in Michigan (E). Such observations suggested that geographic distance may also affect the bacterial community similarity in the biofilters. A statistically significant negative correlation ($P=1.78\times10^{-22}$) was observed between the weighted UniFrac similarity and the geographic distance between the full-scale biofilters (Fig. 4-3), with a z value of 0.00573±0.00058.

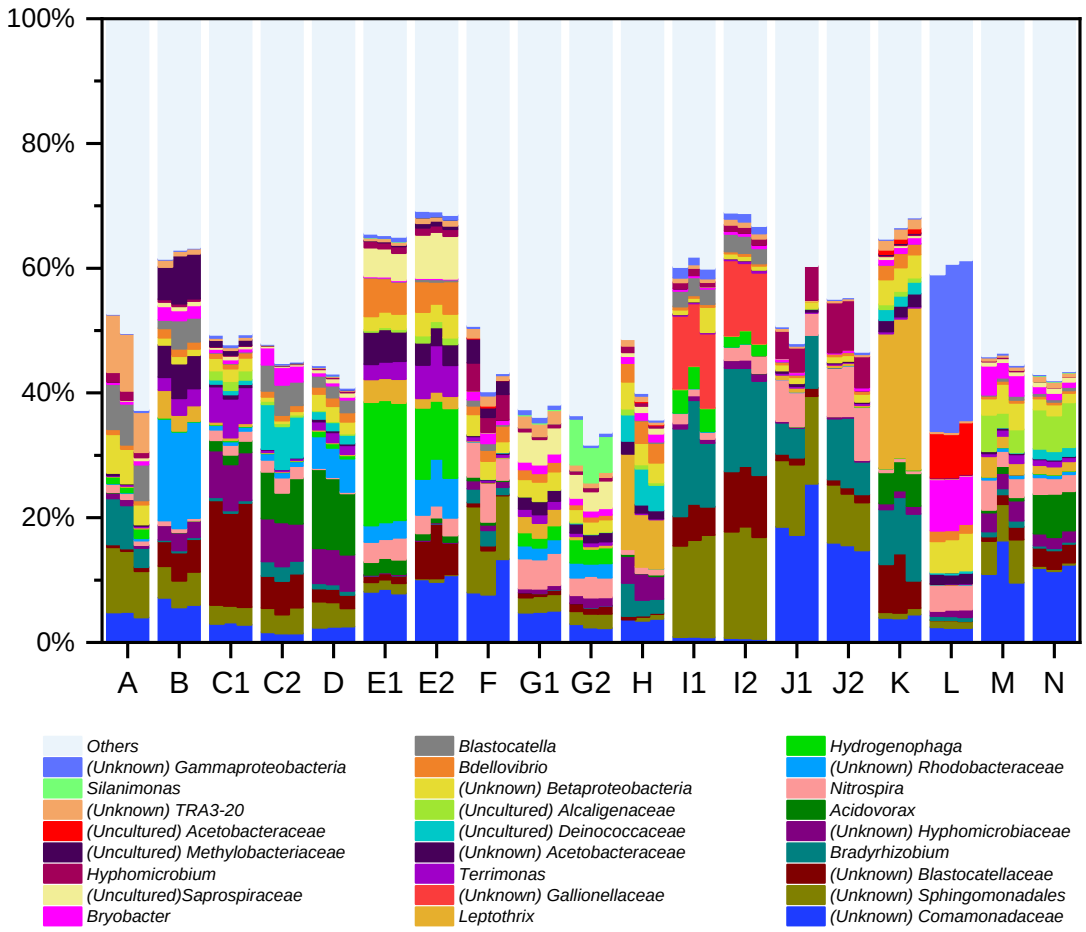

Figure 4- 1. Bacterial community profiles of 26 most prominent genera and unclassified sequences represented in biofilter media samples. The abundant genera are those that comprised greater than 5% of total sequences in at least 3 of 57 samples in total that were collected in this investigation. The genera that are not considered as abundant genera are grouped in “others”. X-axis is the sample ID.

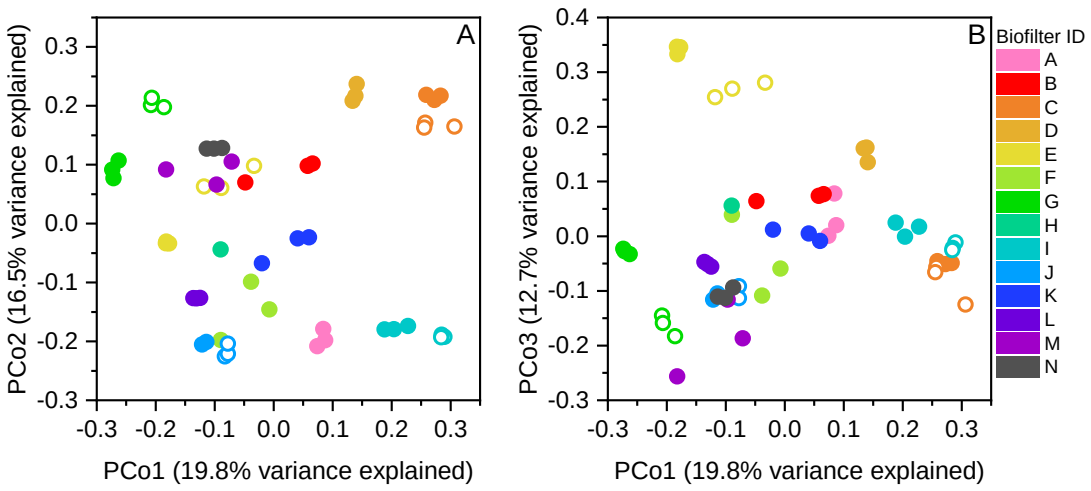

Figure 4- 2. PCoA plots (**A**: PCo1 vs PCo2; **B**: PCo1 vs. PCo3) based on weighted UniFrac distances showing the variations in bacterial communities in the full-scale biofilters. Biofilter IDs are labeled in different color. Open symbol represents the result from a repeating sample collected from the same biofilter three months apart.

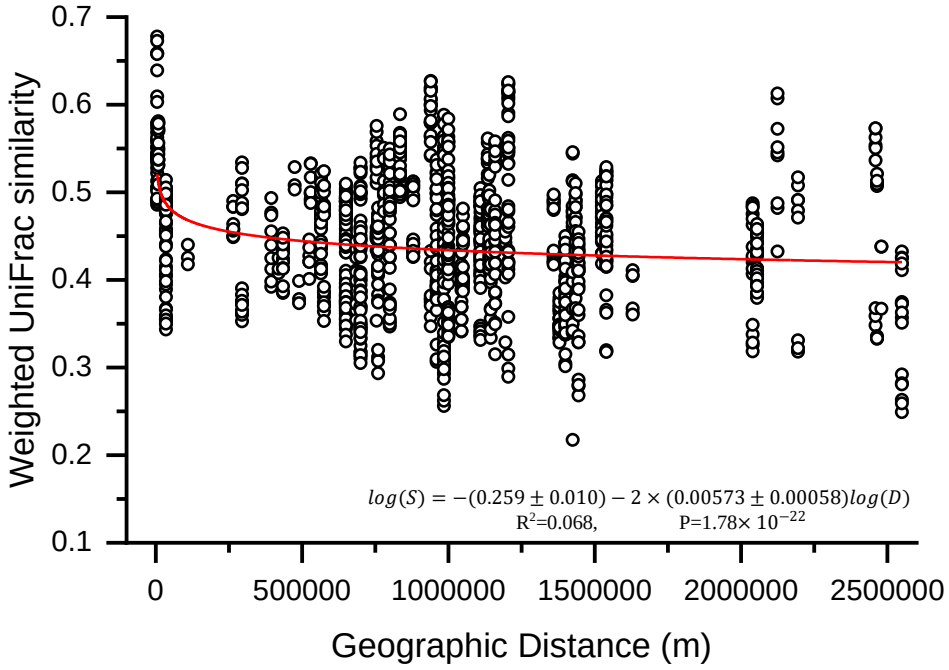

Figure 4- 3. Distance-decay relationship for the bacterial communities in the full-scale biofilters. Each circle represents the pairwise weighted UniFrac similarity of the bacterial communities on two media samples.

PERMANOVA tests indicated that the contemporaneously obtained water quality parameters (i.e., pH, temperature, and influent DOC concentration), media type and upstream ozonation had statistically significant effects on the bacterial community composition ($P_{adonis}<0.05$; Table 4-3). The dispersions of the bacterial communities were significantly different on the GAC and anthracite media and with and without prefiltration ozonation ($P_{betadisper}<0.05$; Table 4-3), which may confound the PERMANOVA results. Among these five parameters, temperature explained most of the variation in bacterial community composition, with R^2_{adonis} of 0.135. Still, 53.0% of the variation in the bacterial community composition was not explained, as indicated by the PERMANOVA results (Table 4-3).

Table 4- 3. Results from PERMANOVA (permutation =999) tests employed to test statistically significant effects of media type, ozonation, and contemporary environmental variables (pH, temperature, and influent DOC concentration) on the bacterial community in biofilters.

	R^2_{adonis}	P_{adonis}	$P_{betadisp}$
Media type (GAC vs. Anthracite)	0.067	0.001	0.001
Ozonation	0.086	0.001	0.040
pH	0.095	0.001	0.167
Temperature	0.135	0.001	0.744
Influent DOC concentration	0.086	0.001	0.175
Residual	0.530		

4.3.3 Multiple regression on matrix (MRM) analysis to identify the relative importance of the factors affecting the bacterial community in the full-scale biofilters.

Two factors had coefficients that were not statistically significant at the 95% confidence level in the MRM analysis: media type (B_5; P=0.373) and ozonation application (B_6; P=0.688). After excluding these insignificant parameters, a final regression relationship was obtained as shown in Eq. 4-2 (permutation=999, p=0.001, R^2 =0.446).

$$\boldsymbol{UniFrac} = 0.533 + 0.053 \times \boldsymbol{D_{std}} + 0.037 \times \boldsymbol{pH} + 0.038 \times \boldsymbol{T_m} + 0.012 \times \boldsymbol{DOC_{inf}} \quad (4-2)$$

A comparison of observed weighted UniFrac dissimilarity values and those computed using Eq. 4-2 is shown in Fig. 4-4. The correlation coefficient (R^2) of 0.446 indicates that 44.6 % of the variance in weighted UniFrac dissimilarities between the bacterial communities in the full-scale biofilters can be explained by this regression relationship.

The coefficients for D_{std}, pH, T_m and DOC_{inf} were all positive, indicating that, as expected, greater discrepancies in the geographic distance or contemporaneous environmental variables increased the dissimilarities of the bacterial communities between the biofilters. The relative importance for a factor of interest is identified by comparing the values of the regression coefficients, which represent the variations in the community dissimilarity per normalized unit of the difference in each factor. Among all abovementioned factors, geographic distance contributed the largest regression coefficient (0.053), with pH, T_m and DOC_{inf} contributing to smaller, but significant coefficients (0.037, 0.038 and 0.012,

respectively). The Mantel test indicated that the geographic distance was significantly correlated with all three water quality parameters ($P_{mantel} < 0.05$; Table D-2).

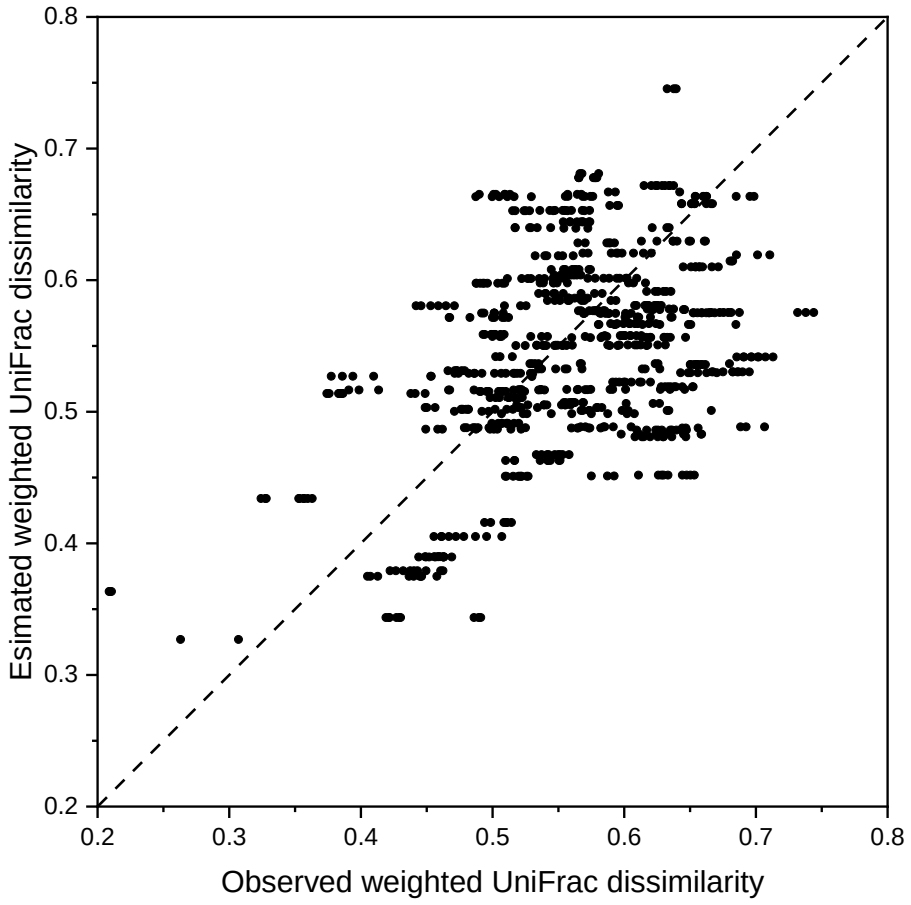

Figure 4- 4. Comparison of the estimated weighted UniFrac dissimilarities computed using empirical relationship developed in the present study to the observed values. The black line represents the 1:1 ratio. The correlation coefficient (R^2) equals to 0.446, meaning 44.6% of the variations in the weighted UniFrac dissimilarities can be explained by using this relationship.

4.4 Discussion

This study provided a comprehensive investigation of the bacterial communities in fourteen full-scale biofilters across North America. The effects of geographic distance, water quality parameters, and engineering decisions on the bacterial community composition were evaluated. Further, the relative importance of each of these factors on the inter-filter variation in community composition was determined using MRM analysis. These findings study should improve our understanding of bacterial geographic patterns in engineered ecosystems.

Bacterial abundance in the biofilters varied with water quality conditions. Positive correlations between the total biomass level and influent DOC concentration were expected because higher DOC concentration should be beneficial for growth of the mesophilic heterotrophic bacteria that tend to dominate the communities in biofilters [78]. The positive correlation with pH is less clear, but may be attributed to enhanced hydrolysis of total organic carbon at higher pH, which increased the assimilable organic carbon level for bacterial growth [111]. Bacterial abundance did not vary significantly over time, which was similar with our previous observations in the pilot-scale biofilters (Chapter 3). Nevertheless, the biofilters were only sampled twice and just a few months apart, so our investigation of temporal (i.e., seasonal) changes in biomass levels is severely limited. It is important to note that the biomass level in the filter media was likely influenced by both contemporaneous as well as historical water quality conditions. The correlation analysis

here, which used only the water quality data measured on the day of sample collection, does not account for the effects of historical water quality conditions. It is expected that the water quality in the few weeks to perhaps few months preceding our sampling may have also impacted the bacterial abundance and community composition in the biofilters.

The biofilter communities shared similarity with other natural and engineered ecosystems. The prominent classified genera, such as *Bradyrhizobium, Nitrospira,* and, *Hydrogenophaga,* are commonly found in freshwater environments [73], soil [74], and water distribution systems [68, 112]. Bacteria from genera that include opportunistic pathogens (e.g., *Legionella* and *Mycobacterium*; [77]) were also detected in the biofilters as minor populations (Fig. D-1), comprising less than 0.1% of the total community. Considering the biofilter microbiome as an integral part of the source-to-tap microbiome, the present of these genera in the biofilters represents a potential risk to drinking water safety and public health.

The *z* value for the distance-decay relationship for the microbiome in the biofilters (0.00573±0.00058) was much lower than *z* values obtained for ecosystems in natural environments, like salt marshes and freshwaters, which were typically greater than 0.02 [103, 105, 109]. This was expected because bacterial communities in biofilters are influenced by the ‘terroir’ or the conditions in the local aquatic environment such as the microbiome in the source water as well as engineering design decisions (i.e., filter media type and pre-filtration treatment). The microbiome in the source water, freshwater

environments such as a lake or river, should exhibit a strong distance-decay relationship as observed for other natural ecosystems. The use of similar pretreatment procedures in drinking water facilities, such as coagulation, flocculation, and sedimentation, as well as similar biofilter operational designs and conditions (media type and backwash strategy; Chapter 3), however, can re-shape the bacterial communities in the source water and reduce the natural biogeographic heterogeneity. Similar results have been observed in other engineered systems, such as activated sludge bioreactors in wastewater treatment plants [107, 108].

Combining statistical analyses like PERMANOVA and MRM, this study provided novel insights into the relative importance of geographic location (i.e., distance), water quality parameters, and engineering design decisions on the bacterial community dissimilarity in biofilters. The contemporaneously measured water quality parameters (i.e., pH, temperature, and DOC concentration) and geographic distance had strong effects on bacterial community composition, which was similar to previous studies on various natural and engineered ecosystems [105, 107, 113]. Although geographic distance had the highest regression coefficients in the MRM analysis, the water quality factors were stronger than the effect of geographic distance when considering all three water quality parameters together (MRM regression coefficient: 0.087 vs. 0.053). This was expected because the geographic distance effect is often diminished in engineered ecosystems, as previously discussed. Furthermore, geographic distance was auto-correlated with the contemporaneous environmental factors (Table D-2), meaning that the coefficient for the

geographic distance in the MRM analysis was likely to be overestimated, which agreed with previous observations [107, 113]. Still, evaluation of the geographic distance effect was valid because the biofilters were isolated ecosystems, meaning the ecological dispersions were likely to be limited [113, 114]. And it also accounted for other unmeasured water quality parameters, historical conditions, and contingencies [105, 107].

Among all three contemporary environmental variables, temperature contributed the most variation in the bacterial communities, which was supported by both PERMANOVA and MRM analyses. Combining all abovementioned factors, still more than 50% variance in the bacterial community was not explained. It is important to note that the analyses here only included the effect of contemporaneous water quality conditions. The unaccounted variance was likely attributed to historical water conditions and pre-filtration treatment procedures, and other unmeasured water quality parameters, such as dissolved oxygen, inorganic nutrients (e.g., phosphorous, nitrogen) and natural organic matter composition [115]. A significant effect of media type and ozonation was indicated by the PERMANOVA results ($P_{adonis}<0.05$) but not by the MRM analysis, which agreed with our hypothesis that the significant different dispersions between different media type or ozonation application confounded the PERMANOVA results. The effect of media type on the bacterial communities was previously observed in bench-scale and pilot-scale studies [15, 116]. More research is needed to evaluate the effects of other engineering designs on bacterial communities in biofilters.

In summary, this study provides a comprehensive understanding of biogeographic patterns of the bacterial communities in full-scale biofilters across North America. The bacterial communities were highly diverse and exhibited inter-filter variations. A significant but weak distance-decay relationship was observed by comparing the weighted UniFrac similarity and the geographic distance between the water treatment facilities. The contemporaneous water quality factors, including pH, temperature, and influent DOC concentration, had a significant and stronger impact on the bacterial communities compared to the geographic distance, even considering that the water quality and geographic distance were autocorrelated. Future research should explore the effects of historical water quality conditions on the bacterial communities in biofilters. Analyses of metagenomes and metatranscriptomes would also be useful for elucidating any geographic patterns in the biological functions of the biofilters.

4.5 Acknowledgements

Financial support for this work was provided by the Water Research Foundation (WRF 4669). This work was carried out in part using Illumina HiSeq instruments at Genomic Center at University of Minnesota, Twin Cities. We thank Jason Carter and Brent Alspach, Carrie Smith, Allison Wheeler of Arcadis, Sarah Page of the City of Ann Arbor, and Caroline Russell of Carollo for helping coordinate this study and providing technical input. We thank all participating drinking water facilities for providing water quality data and treatment information. We also thank Dr. Michael Waak for his inputs in the DNA sequencing analytical methods.

Chapter 5

Multi-scale investigation of nitrifying microorganisms in biofilters

The presence and dynamics of nitrifying microbial community, including ammonia oxidizing bacteria (AOB) and archaea (AOA), and nitrite oxidizing bacteria (NOB), were investigated in full-scale biofilters in fourteen drinking water treatment facilities across North America and in pilot-scale biofilters that were operated at a drinking water treatment facility for 18 months. The abundance of AOA and AOB were quantified by using real-time quantitative polymerase chain reaction (qPCR) targeting ammonia monooxygenase subunit A (*amoA*) gene, and nitrifying bacterial community composition was investigated by sequencing 16S rRNA gene (V3 region) on Illumina HiSeq platform. Our results demonstrated that *Nitrosomonas oligotropha*-like AOB were generally more abundant than AOA in the full-scale biofilters (mean concentration: 5.1±1.5 vs. 3.6±1.2 log *amoA* gene copies/cm^3 bed volume) and in the pilot-scale biofilters (mean concentration: 6.7±0.7 vs. 3.6±0.7 log *amoA* gene copies/cm^3 bed volume). The abundance of *N. oligotropha*-like AOB exhibited temporal variations in the pilot-scale biofilters, which increased with influent ammonia concentration. Although *N. oligotropha*-like AOB were effectively removed (58±44% on average) by the pilot-scale biofilters for most of the operation, shedding from the biofilters periodically served as the primary source of the AOB and largely impacted AOB community composition in the filter effluent primarily when the

AOB concentration in the biofilters was high (> 7 log *amoA*-AOB gene copies/cm^3 bed volume). For samples with a relatively high abundance of uncultured *Nitrosomonadaceae* (i.e., most full-scale biofilter media samples) in the AOB communities, the qPCR assay targeting *amoA*-AOB gene seriously underestimated the abundance of bacteria currently known as AOB. The NOB/AOB ratios exceeded the theoretical ratio for conventional two step nitrification (0.5) in most full-scale biofilters (12 of 14 biofilters) and in the pilot-scale biofilters for most of the operation. This study provides a comprehensive understanding of the nitrifying microbial communities in drinking water treatment biofilters.

5.1 Introduction

Nitrification, the process by which microorganisms oxidize ammonia to nitrate, is an important step in the biogeochemical nitrogen cycle. Conventional nitrification consists of two steps carried out by two different microbial populations: oxidation of ammonia to nitrite by ammonia-oxidizing bacteria (AOB) or archaea (AOA) and oxidation of nitrite to nitrate by nitrite-oxidizing bacteria (NOB). AOB are among the genus *Nitrosococcus* and family *Nitrosomonadaceae*, including the genera *Nitrosomonas* and *Nitrosospira* [117, 118]. Most known AOA belong to the phylum *Thaumarchaeota* with genus level phylogenetic information being very limited [119]. Bacteria of the genera *Nitrospira*, *Nitrolancea, Nitrotoga*, *Nitrococcus*, *Nitrospina* and *Nitrobacter* are capable of nitrite oxidation [120, 121]. Recent studies have shown that some *Nitrospira spp.* are capable of complete nitrification of ammonia to nitrate (i.e., comammox) [122, 123].

Rapid granular media filters inhabited by viable bacteria are termed biologically-active filters or biofilters. Biofiltration is not a new concept, as biofilters operated at low filtration rates, termed slow sand filters, have been used for water treatment since the 1800s [12]. The presence of nitrifying bacteria in water treatment biofilters can affect water quality by not only oxidizing ammonia to nitrite and nitrate but also by degrading organic contaminants, including trihalomethanes, trimethoprim, and estrogenic compounds (e.g., estrone and estriol) [53, 124, 125]. Water treatment facilities that use chloramine as residual disinfectant, however, consider nitrifying microorganisms a potential threat

primarily because they can cause rapid chloramine decay in water distribution systems and associated risks to drinking water safety and public health [16, 17]. Hence, an improved understanding of nitrifying microorganisms in biofilters and their ability to be mobilized from the filter media into the filtrate is needed.

A multi-scale study was performed to investigate factors that affect the abundance and community composition of nitrifying microorganisms in water treatment biofilters. Pilot-scale experiments were performed to assess temporal dynamics and effects of filter operation on nitrifying microbial communities and the export of nitrifiers into the filter effluent. In addition, media was collected from fourteen full-scale biofilters across North America to assess the validity of results from the pilot-scale investigation. The abundance of AOA and AOB were quantified using real-time quantitative polymerase chain reaction (qPCR) targeting ammonia monooxygenase subunit A (*amoA*) gene in archaea and bacteria, respectively, and nitrifying community composition was investigated by sequencing the V3 region of the 16S rRNA gene on Illumina HiSeq platform. The potential biases of qPCR assays for quantifying AOB were evaluated by comparing the qPCR results and DNA sequencing results. Finally, the abundance of NOB and its relationship with AOB were investigated.

5.2 Methods

5.2.1 Sample collection and DNA extraction

Full-scale biofilter media sampling. Filter media samples were collected from fourteen full-scale drinking water treatment biofilters across North America. The biofilters were selected to provide a range of influent water qualities, media types (GAC and anthracite), and pre-filtration treatment processes (e.g., with and without ozonation), as described in Chapter 4 (Table 4-1). Triplicate media samples (approximately 0.5 g as wet weight) were collected in sterile microcentrifuge tubes from the top 7.5 cm of each filter bed between 03/2017 to 06/2017. In total, 42 biofilter media samples were collected for analyzing nitrifying microorganisms. The ammonia concentrations were provided by the participating drinking water treatment facilities (detection limit=0.1 mg N/L).

Pilot-scale biofilter design and media sampling. A pilot-scale biofiltration system was constructed at the Minneapolis Water Treatment and Distribution services (MWTDS) Fridley facility, in Fridley, Minnesota, USA, as described in Chapter 2 (Fig. 2-1). In brief, water from the full-scale recarbonation chambers was used as the feed water for the pilot plant. The water was adjusted to pH 8.5 using hydrochloric acid, coagulated using 2 mg/L (as Fe) ferric chloride and then settled for 1 to 2 hours before entering the biofilter columns. Although eight biofilter columns were operated in parallel for 18 months, only six columns are relevant for this chapter, including three GAC-sand biofilters and three anthracite-sand

biofilters. Media samples (about 0.5 g as wet weight) from four different depths (top, 0.15m., 0.30 m and 0.45m) in all biofilter columns were collected monthly into sterile microcentrifuge tubes. Corresponding pre-and post-filtration water samples (1 L) were also collected in pre-sterilized polypropylene bottles. In total, 96 media samples and 117 water samples were collected from the pilot-scale biofilters. The water temperature was measured daily during weekdays. The ammonia in water samples were determined weekly according to EPA method 350.1 from 09/2015 to 01/2017. The nitrite and nitrate concentrations in water samples were determined weekly according to EPA method 353.2 from 09/2015 to 11/2015 and 04/2016 to 11/2016.

Sample preparation and DNA extraction. All water and media samples collected from the pilot-scale biofilters were transported on ice to the laboratory. All media samples collected from the full-scale biofilters were frozen and then shipped on ice to the laboratory via overnight carrier. The water samples were filtered through pre-sterilized nitrocellulose filters (diameter = 47 mm; pore size = 0.2 μm) within 2 hours of the sample collection. The nitrocellulose filters and media samples were then submerged in 0.6 mL lysis buffer (120 mM sodium phosphate, 5% SDS, pH 8.0), subjected to three consecutive freeze-thaw cycles and incubated at 70 ^{o}C for 90 minutes. Then, genomic DNA in each sample was extracted using the FastDNA SPIN Kit for Soil (MP Biomedicals, Solon, OH). Purified DNA was stored at -20°C until future use. After extraction, the media samples in the microcentrifuge tubes were removed, dried at 105 ^{o}C, and then weighed to obtain the dry weights of the media. The corresponding media volumes were computed from the dry

weights using typical media bulk density values and then the volumes were used for normalization of the qPCR results.

5.2.2 Real-time quantitative polymerase chain reaction (qPCR)

Real-time quantitative qPCR was used to detect and quantify three genes of interest [126–128]: The 16S rRNA gene for *Bacteria* that was used to quantify total bacterial biomass, and the *amoA* gene for AOA (*amoA*-AOA gene) and AOB (*amoA*-AOB gene) that was used to quantify AOA and *Nitrosomonas oligotropha*-like AOB. All media samples were subjected to qPCR measurement using the protocols of LaPara et al., (2015)[13]. Standards for qPCR were prepared from serial dilution of synthetic double-stranded DNA fragments (gBlocks) to develop calibration curves for quantification of gene copies in each sample. The censoring limit (CL) was defined as the lowest standard of a given qPCR assay that amplified (1100 copies per reaction for 16S rRNA gene, 6.33 copies per reaction for *amoA* gene), normalized by volume (L) of the water sample and the bed volume (cm^3) of the media sample used for DNA extraction. The detailed information of the gene targets and respective primer sequences and gBlock standards is shown in Table C-1 and E-1. A two-sample t-test was used to determine if there was a significant difference between *amoA*-AOB gene and *amoA*-AOA gene concentrations, and if there was a significant difference in gene concentrations in different samples collected from the pilot-scale biofilters on each sampling dates, and the p-values were adjusted using Bonferroni correction [70]. Pearson's

correlation tests were conducted to determine if a statistically significant correlation existed between two results (e.g., gene concentration and ammonia concentration).

5.2.3 High-throughput 16S rRNA gene sequencing

The DNA extract from each sample was amplified using PCR, purified, quantified, and then pooled by equal mass to create a library for sequencing, as previously described [13]. Next generation high-throughput sequencing of PCR-amplified 16S rRNA gene fragments (V3 region) was performed on the Illumina HiSeq platform at the University of Minnesota Genomics Center (UMGC) using primers 341F and 534R [63]. Unprocessed sequence reads have been uploaded to Sequence Read Archive (accession: PRJNA521294 (full-scale biofilter media) and PRJNA515121 (pilot-scale biofilter media)).

The sequence reads were demultiplexed, trimmed, and filtered using QIIME2 (version 2018.2). Amplicon sequence variants (ASVs) were determined using 'DADA2' [64, 65] and then assigned consensus taxonomy using SILVA rRNA database (release 128) [66]. The concentrations of *Nitrosomonas*-like, AOB-like and NOB-like ASVs were determined by multiplying their respective relative abundance values (i.e., the ratio of the number of respective ASVs detected to the total number of sequence reads) from the DNA sequencing results by the 16S rRNA gene concentrations measured via qPCR. Paired t-test was used to determine if there was a significant difference in gene concentrations between the pilot-scale GAC-sand and anthracite-sand biofilters over the entire experiment. Bray-Curtis

dissimilarity in AOB communities between samples were calculated using ‘vegan’ package after sequence libraries were randomly trimmed down to 1,000 sequences per profile. Principal coordinates analyses (PCoA) plots were made to visualize the Bray-Curtis dissimilarity using *cmdscale* command in R.

5.3 Results

5.3.1 Concentration of ammonia, nitrite and nitrate in filter influent and effluent

In the full-scale investigation, influent ammonia concentrations were lower than the detection limit (0.1 mg N/L) for all biofilters on the day of sample collection. In the pilot-scale investigation, water temperature ranged from 3 to 28 ^{o}C. Ammonia concentrations in the filter influent ranged from below detection limit (0.015 mg N/L) to 0.131 mg N/L, with a mean of 0.048±0.029 mg N/L (Fig. E-1-A). The mean removal of ammonia in pilot-scale GAC-sand and anthracite-sand biofilters were 68.6±24.1% and 50.8±37.0% (Fig. E-2). No significant difference (P=0.41; Paired t-test) was observed between the GAC-sand and anthracite-sand biofilters for most of the operation, except during November 2015 to March 2016, when a significantly higher ammonia removal was observed in the GAC-sand biofilters ($P=1.9\times10^{-4}$; Paired t-test). Nitrite concentrations in most pre-filtration water samples (32 of 36 samples) collected from the pilot-scale biofilters were below the detection limit (0.015 mg N/L; Fig. E-1-B). Nitrate concentration in the filter influent were relatively stable over time (mean±S.D.: 0.97±0.51 mg N/L), with a periodically spiking concentration (up to 2.28 mg N/L) observed in June 2016 (Fig. E-1-C).

5.3.2 Quantification of *amoA* genes in the biofilters

The concentrations of *amoA*-AOB and *amoA*-AOA genes varied from 3.3±0.3 to 7.2±0.1 log copies/cm^3 bed volume and below the CL to 6.5±0.1 log copies/cm^3 bed volume, respectively (Fig. 5-1). Media from biofilters E, G and J had high concentrations of *N. oligotropha*-like AOB but low concentrations of AOA. Media from biofilter I was unusual in that it was the only sample with significantly higher AOA than AOB (5.9±0.2 vs. 5.2±0.1 log *amoA* gene copies /cm^3 bed volume; P=0.006). Media samples from some facilities (e.g., C, D, H, and L) had relatively low levels of both AOB and AOA, suggesting minimal capacity for microbial nitrification. The normalized results (i.e., the ratio of *amoA* genes to 16S rRNA genes) show that while ammonia oxidizers can account for as much as 1% of biofilter communities, they typically represent 0.01% or less (10 of 14 biofilters; Fig. E-3).

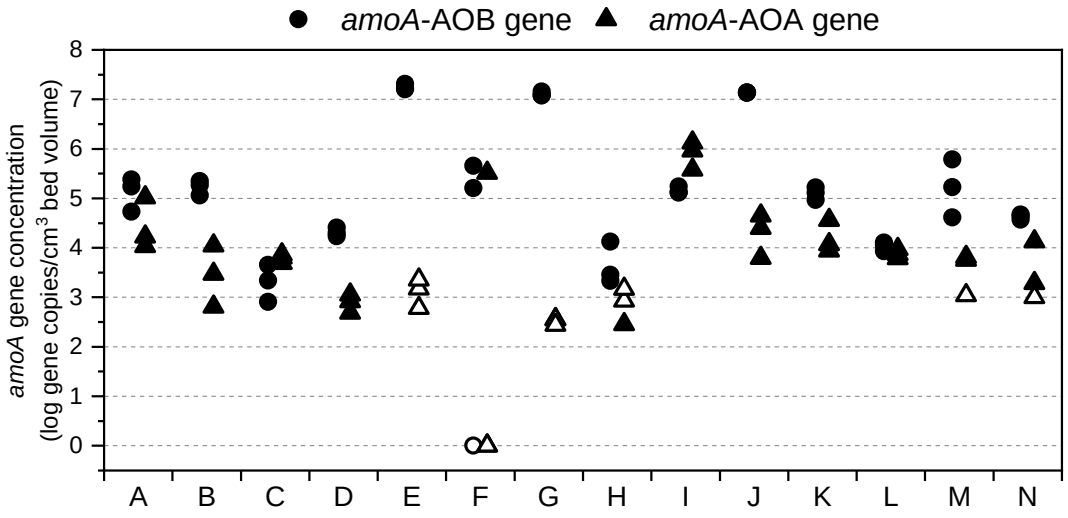

Figure 5- 1. Concentrations of *amoA*-AOB gene and *amoA*-AOA gene on full-scale biofilter media samples. X-axis is the biofilter ID. Open symbols indicate that the results are below the censoring limit.

The AOB concentration (i.e., *amoA*-AOB gene concentration) on the media samples collected from the top of the pilot-scale GAC-sand and anthracite-sand biofilters ranged from 5.4±0.2 to 7.8±0.2 and 5.2±0.1 to 7.6±0.4 log copies/cm^3 bed volume, respectively, comprising on average 1.2±1.6% and 1.7±3.7% of the total biomass, respectively (Fig. 5-2 and E-4). No significant difference (P_{adj}>0.05) in the *amoA*-AOB gene concentrations on the GAC and anthracite media samples was observed on most sampling dates, except on 03/30/2016, when a significantly higher concentration was observed on the GAC media (P_{adj}=0.01). The AOB concentrations were positively correlated with the ammonia concentration in the filter influent in the GAC-sand and anthracite-sand biofilters (Pearson's correlation: 0.53 (P=0.05) and 0.68 (P=0.01), respectively) but negatively correlated with the water temperature (Pearson's correlation: -0.63 (P=0.02) and -0.52 (P=0.05), respectively). The mean concentrations of *amoA*-AOB genes in the media samples decreased with depth in the biofilters (Fig. E-5), with a mean reduction of 0.6±0.6 and 0.7±0.7 log copies/cm^3 bed volume for every 0.15 m of depth increase in the GAC-sand and anthracite-sand biofilters, respectively. The *amoA*-AOA gene concentrations were much lower than the *amoA*-AOB gene concentrations (approximately 2-3 log copies lower in gene concentrations) in the respective samples (Fig. E-6).

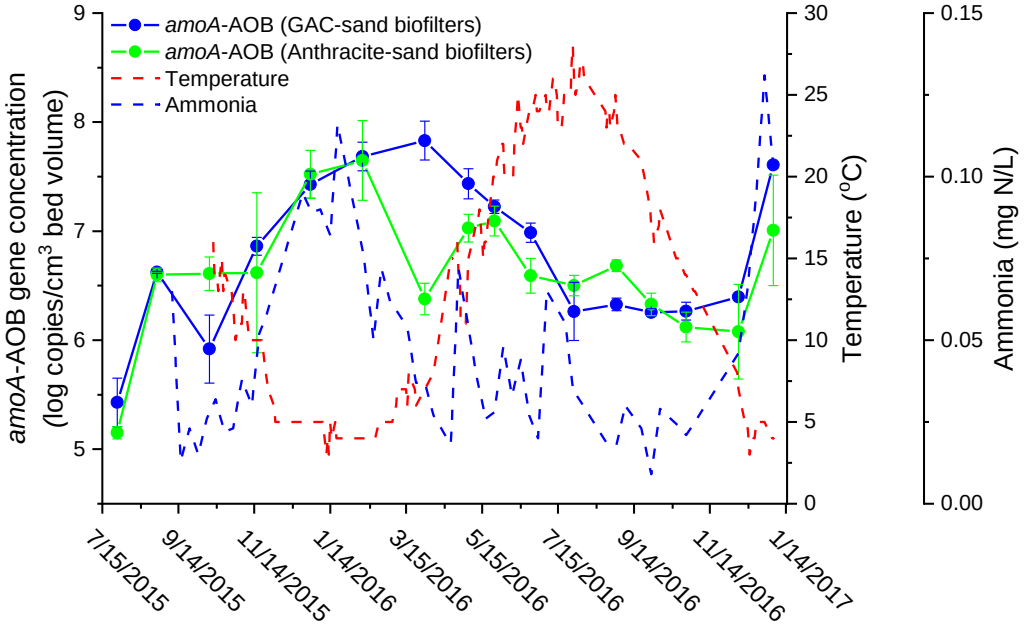

Figure 5- 2. *amoA*-AOB gene concentrations on media samples collected from the pilot-scale GAC-sand (blue solid line) and anthracite-sand (green solid line) biofilters. Red dashed line represents the water temperature. Black dashed line represents the influent ammonia concentration. All gene concentrations are above censoring limits. Error bars represent the standard deviation computed from triplicate samples.

amoA-AOB genes were also detected in the filter influent, and filter effluent from GAC-sand and anthracite-sand biofilters, ranging from 3.3±0.4 to 5.1±0.1, 2.7±0.3 to 5.2±0.1, and 3.0±0.2 to 5.2±0.1 log copies/L, respectively (Fig. 5-3). The GAC-sand and anthracite-sand biofilters effectively removed AOB (64±22% and 48±56% on average, respectively) from the filter influent on most sampling dates, expect on a few dates in the winter (e.g., 02/09/2016 and 01/04/2017), when higher concentrations of *amoA*-AOB gene were observed in the filter effluent compared to the influent (Fig. 5-3).

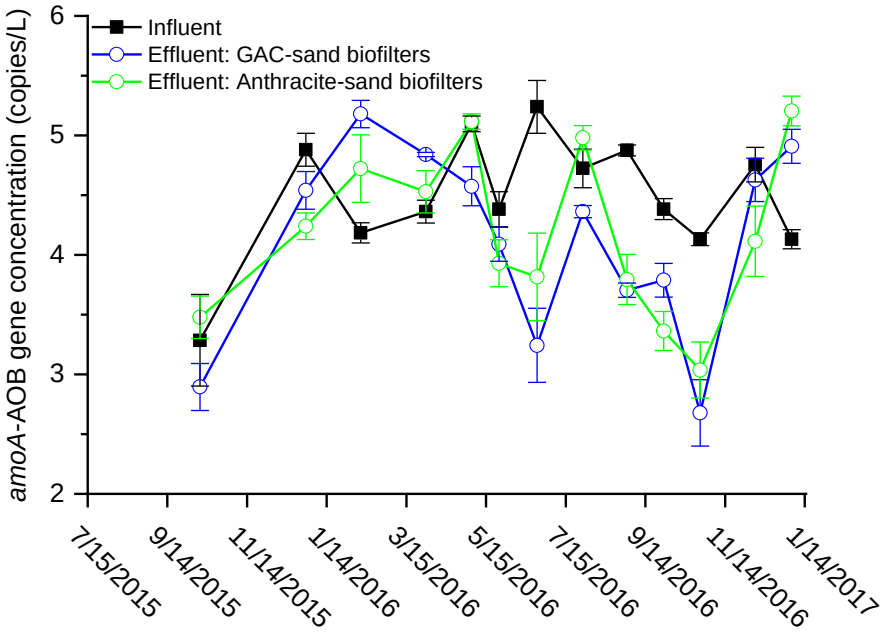

Figure 5- 3. *amoA*-AOB gene concentration in the filter influent and filter effluent from the pilot-scale GAC-sand and anthracite-sand biofilters. Error bars represent the standard deviation computed from triplicate samples.

5.3.3 Ammonia oxidizing bacterial communities in the biofilters

AOB-like ASVs from family *Nitrosomonadaceae* (612 different ASVs) were consistently detected in the full-scale and pilot-scale drinking water treatment biofilters according to the DNA sequencing results (Fig E-7). Other AOB, such as genera *Nitrosococcus*, were not detected. Genus *Nitrosomonas* and uncultured population in family *Nitrosomonadaceae* were the most prominent population in AOB, comprising a total of 97.5±4.6 % of the communities. The concentration of *Nitrosomonas*-like ASVs in the pilot-scale biofilter media samples increased with ammonia concentration and decreased with temperature (Fig. E-8), which agreed with the *amoA*-AOB gene concentration trends (Fig.

5-2). The abundance of uncultured *Nitrosomonadaceae*, however, was relatively stable throughout the entire course of the experiment, showing no significant correlation with influent ammonia concentration ($P>0.05$; Pearson's correlation; Fig. E-8). Interestingly, the abundance of uncultured *Nitrosomonadaceae* appeared to be affected by media type, as concentrations on the GAC media were significantly higher than those on the anthracite media ($P=1.3\times10^{-5}$; paired t-test).

PCoA plots were prepared to visualize the difference in AOB communities between filter influent, biofilter media and filter effluent in the pilot-scale GAC-sand and anthracite-sand biofilters (Fig. 5-4). The separation between water samples and media samples in the PCoA plots suggests that the AOB communities in the water samples and the media samples were different. Furthermore, the close proximity of filter influent and effluent suggested that those communities were similar. Filter effluent was relatively closer to biofilter media compared to filter influent in the PCoA plot, especially when the water was cold (e.g., 02/09/2016 in Fig. 5-4). Bray-Curtis similarities of AOB communities between filter influent, biofilter media, and filter effluent were also compared (Fig. E-9). Relatively higher similarity values were observed between the biofilter media and filter effluent than between the media and filter influent. The similarity values between the filter influent and effluent were higher than those between media and water on most sampling dates, except on a few dates in the winter (e.g., 02/09/2016 and 01/04/2017), when higher similarity values were observed between biofilter media and filter effluent (Fig. E-9).

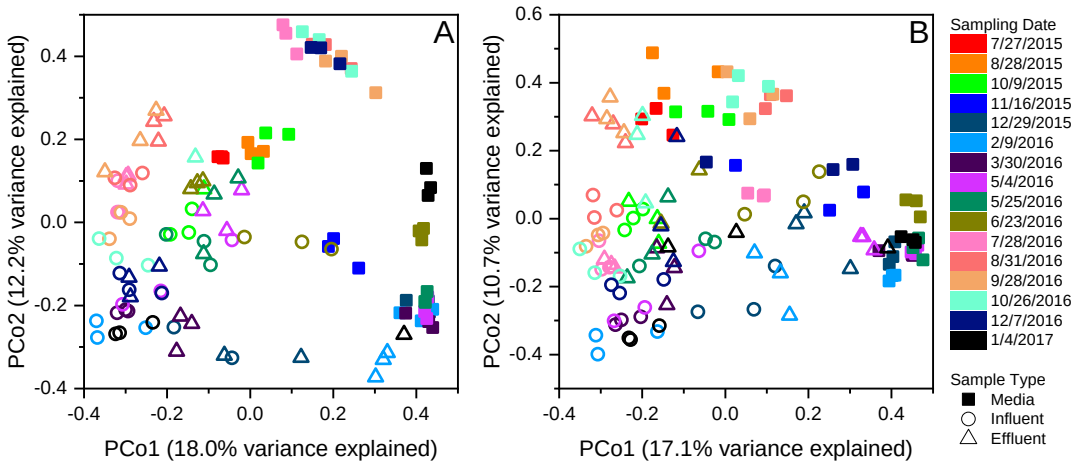

Figure 5- 4. PCoA plots based on Bray-Curtis dissimilarity showing the variations in the ammonia oxidizing bacterial communities in the filter influent, biofilter media and filter effluent collected from the pilot-scale GAC-sand (**A**) and anthracite-sand (**B**) biofilters. Different color represents the sampling date in the pilot-scale investigation.

The theoretical ratio of *amoA*-AOB gene to 16S rRNA gene ranged from 0.13 to 3.0, considering that AOB contain 2-3 copies of the *amoA*-AOB gene and 1-15 copies of the 16S rRNA gene per genome [118, 127, 129–131]. The ratios of *amoA*-AOB genes to *Nitrosomonas*-like ASV ranged from 0.04 to 14.2 (mean of 0.91) and were within the theoretical range for 95 out of 138 media samples collected from the full-scale and pilot-scale biofilters (Fig. 5-4-A). The ratios of *amoA*-AOB genes to AOB-like ASV in the pilot-scale biofilter media (range: 0.07 to 5.36; mean: 1.29) also generally fell within the theoretical range (Fig. 5-4-B). The ratios of *amoA*-AOB genes to AOB-like ASV for the full-scale biofilters (range: 5.4×10^{-7} to 0.20; mean: 0.03), however, generally were below the lower limit of the theoretical ratio.

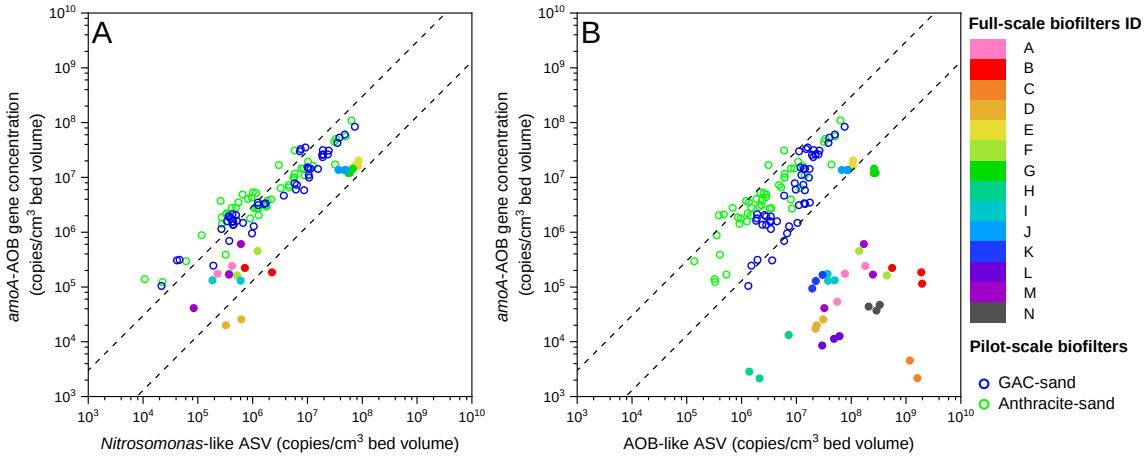

Figure 5- 5. Comparison of *amoA*-AOB gene concentration measured using qPCR and abundance of *Nitrosomonas*-like ASV (**A**) and AOB-like ASV (**B**) according to DNA sequencing results. Region between dashed lines represents the range of theoretical ratio of *amoA*-AOB gene to 16S rRNA gene in AOB.

5.3.4 Comparison of NOB and AOB in the biofilters

ASVs from the genera *Nitrospira*, *Nitrolancea*, and *Candidatus Nitrotoga*, were the only known NOB detected in the full-scale and pilot-scale biofilters (78 different ASVs), with *Nitrospira* being the most abundant, representing 100.0±0.3% (mean ± standard deviation) of the NOB. The NOB-like ASV concentrations in the full-scale biofilters ranged from 5.3±0.1 to 7.8±0.4 log copies/cm^3 bed volume, comprising on average 2.5±1.9% of the total community (Fig. E-10-A). The NOB-like ASV concentrations were relatively stable in the pilot-scale GAC-sand biofilters after the first 2 months of operation, ranging from 7.1±0.0 to 8.0±0.2 log copies/cm^3 bed volume, comprising on average 3.1±2.1% of the total community (Fig. E-10-B). The NOB-like ASV concentrations in the pilot-scale anthracite-sand biofilters were similar to those in the GAC-sand biofilters for most of the

18-month operating period, except during 03/2016 to 08/2016, when significantly lower concentrations were observed in the anthracite-sand biofilters (Fig. E-10-B).

The ratio of NOB to AOB in the full-scale biofilter media samples ranged from -0.7±0.2 to 0.4±0.0 (log_{10} mean±S.D.; Fig. 5-6-A). The values for all media samples, except those from Biofilter B and C, exceeded the theoretical ratio of -0.3 (i.e., $10^{-0.3}$ = 0.5) for conventional two step nitrification [132, 133]. Despite variations in the NOB/AOB ratio, the NOB-like ASV concentration correlated positively with the AOB-like ASV concentration (Pearson's correlation: 0.49; P=0.00) in the full-scale biofilters. The NOB/AOB ratio values ranged from -0.9±0.2 to 1.1±0.2 (log_{10} mean±S.D.) in the pilot-scale GAC-sand biofilters and exhibited temporal variations (Fig. 5-6-B), increasing with temperature and decreasing with influent ammonia concentration (Pearson's correlation: 0.6 (P=0.03) and -0.6 (P=0.03), respectively). A similar trend was observed in the pilot-scale anthracite-sand biofilters (Fig. 5-6-B), but the correlation coefficients with temperature and ammonia concentration were not significantly different from zero at the 95% confidence level (P=0.60 and 0.18, respectively). The NOB/AOB ratios in the pilot-scale biofilters exceeded the theoretical ratio for most of the operation, except during 02/2016 to 05/2016 in the anthracite-sand biofilters.

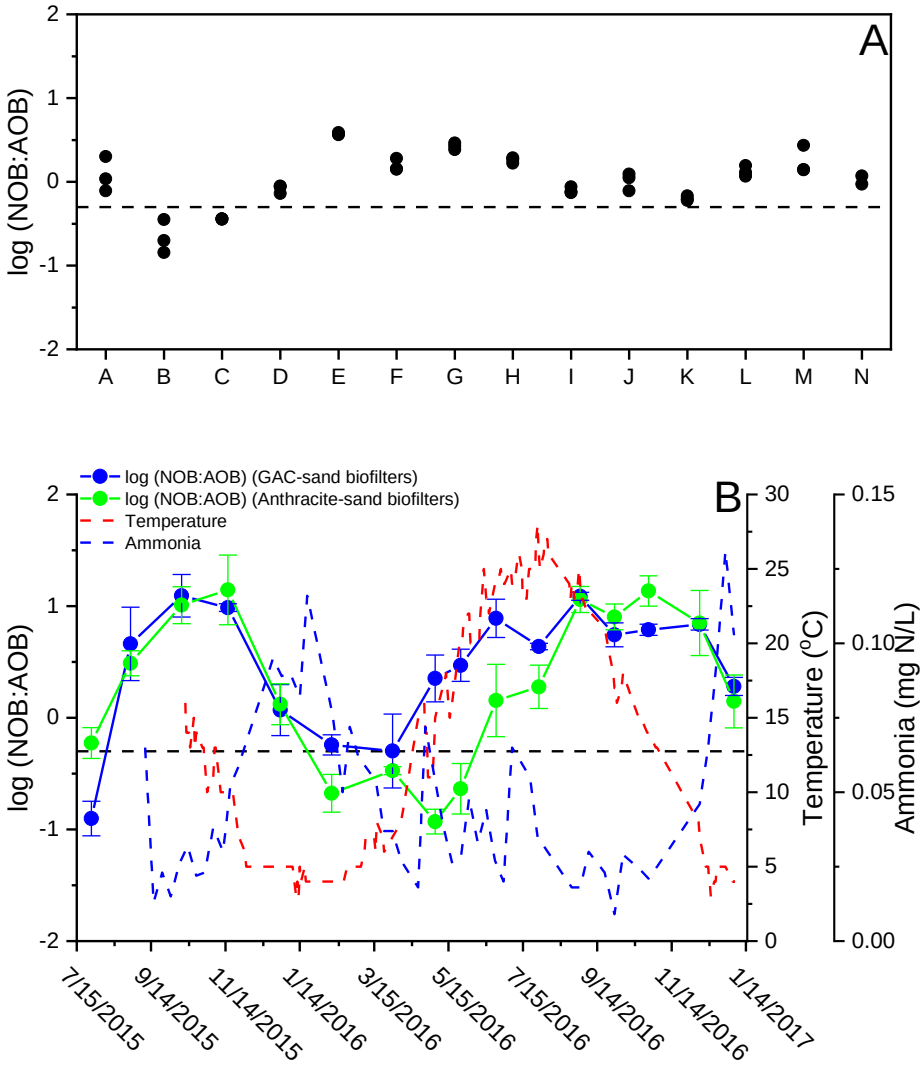

Figure 5- 6. Ratio of NOB to AOB for the full-scale biofilter media samples (**A**) and pilot-scale biofilter media samples over time (**B**) according to DNA sequencing results. Black dashed line represents the theoretical ratio of NOB to AOB (0.5) for a conventional two-step nitrification. X-axis in **A** is the biofilter ID. Error bars in **B** represent the standard deviation computed from triplicate samples.

5.4 Discussion

A unique aspect of this study was the combination of a full-scale biofilter media sampling campaign and an 18-month pilot-scale experiment involving monthly sampling of the filter media and water. The pilot-scale experiment was suitable to investigate the variations and dynamics of nitrifying communities in biofilters, but it may not fully represent the conditions of full-scale biofilters as the conditions, such as water qualities and treatment procedures, varied from system to system. Our results demonstrated that although results from the pilot-scale investigation were mostly similar to the full-scale investigation (e.g., the concentrations of AOB and AOA, and the NOB/AOB ratios), there were some discrepancies between the investigations at different scales. For examples, an agreement between *amoA*-AOB gene abundance and *Nitrosomonadaceae*-like ASV concentration was observed in our pilot-scale investigation, but not in the full-scale investigation (Fig 5-5-B). Interpretation of the results based on the pilot-scale study alone may lead to biased conclusion about the viability of the qPCR assays targeting *amoA*-AOB genes in quantifying AOB. The complementarity of full-scale and pilot-scale investigation in this study ensured a comprehensive understanding of nitrifying microbial communities in biofilters.

In general, AOB was more abundant than AOA in most full-scale biofilters and the pilot-scale biofilters, according to the qPCR results targeting *amoA* genes. The dominance of AOA and AOB in ammonia oxidizing microbial communities in the biofilters were

controversial according to previous studies. [13, 134–137]. Theoretically, AOA should have an advantage in the biofilters, considering that surface waters normally contain very low concentrations of ammonia and AOA have a much higher affinity for ammonia than AOB [138]. One possible explanation for the predominance of AOB in the biofilters is that AOA are believed to form weaker biofilms than AOB, hence, AOA are more likely to be washed out from the biofilter media during backwashing [134, 137].

Temporal variations in AOB concentration in the pilot-scale biofilters were related to influent water quality and mainly the influent ammonia concentration. Previous studies suggested that AOB growth correlated with increasing temperature (12 to 22 ^{o}C) and increasing free ammonia concentration [139, 140]. In our pilot-scale system, AOB concentration on the media positively correlated with influent ammonia but negatively correlated water temperature, as the highest ammonia concentrations were observed in the winter months. Despite the higher AOB concentration in the biofilters, the nitrifying bioactivity was still relatively low in the winter, as indicated by the ammonia removal (Fig. E-2). In fact, the highest ammonia removal was normally observed in the early spring (March to June) and later summer (September to November) in our system, when there was moderate amount of AOB in the biofilters and the temperature was still high enough to keep them activated (Fig. E-2). This result suggests that temperature only matters when there is sufficient ammonia available to support the growth of nitrifiers. Higher concentrations of AOB were observed in the pilot-scale GAC-sand biofilters than anthracite-sand biofilters at the end of winter 2016 (e.g., 03/30/2016). This was probably

due to the catalytical reduction of chloramine in the backwash water (i.e., tap water, contained 2-4 mg Cl_2/L chloramine) by the GAC media, which provided ammonia for AOB growth when influent ammonia concentration decreased in the late winter.

Although AOB in filter influent were effectively removed by the pilot-scale biofilters for most of the operation, shedding from the biofilters periodically served as the primary source of the AOB in the filter effluent. Substantially higher concentrations of AOB in the filter effluent compared to influent were only observed in the winter months (i.e., January to March), when the AOB was doing well in the biofilters (Fig. 5-3). At the same time, the AOB communities in the filter effluent were largely impacted by the biofilters (Fig. 5-4). The biofilters also marginally impacted the ammonia oxidizing bacterial communities in the filter effluent during other period of the operation (Fig. 5-4 and E-9) Taken as a whole, these results suggest that the biofilters were removing and shedding AOB at the same time, and the impact of biofilter shedding was related to the AOB abundance in the biofilters. Previous studies suggested that AOB in the distribution system can cause disinfectant (i.e., chloramine) decay and bacterial regrowth [16, 17]. It is unclear, however, whether the AOB in the filter effluent are able to survive disinfection and then enter the distribution system.

The comparison between the qPCR and DNA sequencing results provided valuable assessment of the potential biases in qPCR-based assays. Our results suggest that even though the qPCR results of *amoA*-AOB genes provided good assessment of *Nitrosomonas*, this assay may seriously underestimate the abundance of bacteria currently known as AOB,

especially for samples with a relatively high abundance of uncultured *Nitrosomonadaceae* (i.e., most full-scale biofilter media samples; Fig. E-7). One possible explanation is that the sequences of the *amoA* genes in uncultured *Nitrosomonadaceae* differ from those in *N. oligotropha*. This would not be surprising as the primers were designed to target the *amoA* gene in *N. oligotropha* [127]. Another possibility is that these uncultured *Nitrosomonadaceae* do not possess the capability for ammonia oxidation (i.e., no *amoA* genes), even though they are phylogenically similar to well-known AOB (e.g., *Nitrosomonas*). The latter explanation is further supported by the lack of a correlation between uncultured *Nitrosomonadaceae* abundance and biofilter influent ammonia concentration in the pilot-scale experiments. It's important to note that the sequencing-based quantitative results in this study could also be biased by differences in DNA extraction efficiency, sequencing depth, and/or sequencing accuracy between samples [141, 142]. More research is needed to determine the viability of the *amoA* gene qPCR-based assay in quantifying AOB and the roles of these uncultured *Nitrosomonadaceae* in microbial nitrification in drinking water treatment biofilters.

The NOB/AOB ratios generally exceeded the theoretical ratio for a conventional two step nitrification, which agreed to many previous observations in nitrifying bioreactors and biofilters [13, 127, 143, 144]. One possible explanation is that AOA in the biofilters were oxidizing ammonia and providing some of the nitrite for the NOB. This was likely not the case because AOA were generally in much lower abundance than AOB in most full-scale biofilters and in the pilot-scale biofilters. Another possibility is that some of the NOB in

the biofilters were comammox *Nitrospira spp.* [122, 123]. Also, *Nitrospira* may not be strictly autotrophic as some can utilize simple organic substrates, such as pyruvate and formate, for carbon assimilation and potential as energy sources for growth and other activities [145, 146]. Thus, at least some NOB in biofilters may not rely on ammonia oxidizers to provide nitrite as electron donor. This is also supported by investigating the NOB abundance in our pilot-scale investigation (Fig. E-10), which did not correlate with influent ammonia concentration. Recent studies also demonstrated that the NOB/AOB ratios were lower than the theoretical ratio in chloraminated drinking water distribution systems due to the oxidation of nitrite by chloramine [112], suggesting that the presence of oxidative chemicals may also affect the NOB/AOB ratios. Regardless, additional research is needed to explain the relationships and roles of AOB and NOB in microbial nitrification.

In conclusion, this study provides a comprehensive understanding of nitrifying microorganisms in biofilters. *N. oligotropha*-like AOB were generally more abundant than AOA in the biofilters and increased with influent ammonia concentration. Biofilters were capable of removing and shedding AOB at the same time. Shedding from the biofilters served as the primary source of the AOB and largely impacted AOB community composition in the filter effluent primarily when the AOB abundance was high in the biofilters. For samples with a relatively high abundance of uncultured *Nitrosomonadaceae* (i.e., most full-scale biofilter media samples), the qPCR assay targeting *amoA*-AOB gene seriously underestimated the abundance of bacteria currently known as AOB. The NOB/AOB ratios in most biofilters exceeded the theoretical ratio of 0.5, which was likely

to due to presence of comammox bacteria and the mixotrophic lifestyle of some NOB. More research is needed to determine the roles of uncultured *Nitrosomonadaceae* and comammox *Nitrospira spp.* in microbial nitrification in water treatment biofilters.

5.5 Acknowledgements

Financial support for this work was provided by the city of Minneapolis and Water Research Foundation (WRF 4669). This work was carried out in part using Illumina HiSeq instruments at Genomic Center at University of Minnesota, Twin Cities. We thank Annika Bankston, George Kraynick and Dr. Li Zhang of Minneapolis Water Treatment and Distribution Services for providing technical input in the pilot-scale experiment. We thank Ashley Evans, Jason Carter and Brent Alspach, Carrie Smith, Allison Wheeler of Arcadis, Sarah Page of the City of Ann Arbor, and Caroline Russell of Carollo for helping coordinate this study and providing technical input. We thank all participating drinking water facilities for assisting the sample collection. We also thank Dr. Michael Waak for his inputs in the DNA sequencing analytical methods.

Chapter 6

Conclusions

The results in this work provide a comprehensive understanding of CECs removal and the microbiome in water treatment biofilters. The primary conclusions of this work are summarized in seven points:

GAC-sand biofilters provided better CECs removal than conventional anthracite-sand biofilters due to dual function of adsorption and biodegradation, and adsorption was the dominant removal mechanism for most CECs [Chapter 2]. In conventional anthracite-sand biofilters, effective CECs removal (>50%) was only achieved for highly biodegradable CECs, like fluoxetine, and primarily at high temperatures (>20 ^{o}C). Thus, compound biodegradability and seasonal temperature changes will determine if and when conventional anthracite-sand biofilters are able to remove CECs. The GAC-sand biofilters, however, serve as a long-term, sustainable, and effective solution of treating multiple CECs simultaneously in drinking water treatment. Moderate to nearly complete CECs removal (mean removal ranged from 49.1±27.7% to 94.4±7.0% for each CEC) was observed in GAC-sand biofilters over a wide range of operational conditions. For most of the CECs tested in this work, adsorption to the GAC media was the dominant removal mechanism in the GAC-sand biofilters. Biodegradation can extend GAC bed life by reducing contaminant loading to sorption sites and regenerating occupied sites.

Operational conditions (hydraulic loading rate and throughput), water quality (temperature, and influent CECs and DOC concentration), and CECs properties (biodegradability and charge) significantly impact CECs removal in GAC-sand biofilters [Chapter 2]. The effects of these factors are explained by considering the dual function of adsorption and biodegradation in GAC-sand biofilters. The adsorption of CECs to GAC media benefits from a greater contact time (low loading rate), higher sorption driving force (high CECs concentration), and an enhanced electrostatic interaction between the positively charged CECs and negatively charged GAC surface. The adsorption capacities, however, decreased with throughput due to the usage of sorption sites. CECs had different biodegradabilities, and the biodegradation of CECs benefits from a greater contact time and higher bioactivity in the biofilters (higher temperature and influent DOC concentration). A multiple linear regression based empirical relationship considering all abovementioned parameters was developed in this work to predict CECs removal in the GAC-sand biofilters, showing good agreement with the experimental data. Caution must be exercised in applying such empirical models because they tend to be site-specific. These findings lead to several recommendations for drinking water treatment facilities in their biofilter operation, including maintaining as low a hydraulic loading rate as possible and periodic GAC media replacement or regeneration for better CECs removal. Also, additional treatment, such as advanced oxidation, may be needed to effectively remove non-biodegradable, negatively charged CECs (e.g., sulfamethoxazole).

Bacterial communities in biofilters were highly diverse, and shared similarity with other natural and engineered ecosystems [Chapter 3 and 4]. Bacterial communities in the biofilters were highly diverse (Shannon Index>5.3), regardless the source water quality and pre-filtration treatment procedures. The prominent classified genera in the biofilter microbiome, such as *Limnohabitans*, *Flavobacterium*, *Nitrospira*, and, *Hydrogenophaga*, are commonly found in freshwater environments, soil, and water distribution systems [68, 73, 74, 112]. Human gut bacteria (e.g., *Bacteroides* [76]) and bacteria from genera that include opportunistic pathogens (e.g., *Legionella* and *Mycobacterium* [77]) were also detected in the biofilters as minor populations (<0.2% of the total community). As a part of the source-to-tap microbiome, the present of these bacteria in the biofilters may pose potential risks to drinking water safety and public health. More testing would be needed to determine whether these risky bacteria are able to survive subsequent disinfection and enter the distribution system.

Water quality (temperature, pH, influent DOC, and ammonia concentration), bed depth, and engineering decisions (media type and backwash strategy) significantly impacted the microbiome in biofilters [Chapter 3, 4, and 5]. In the pilot-scale biofilters, the bacterial community exhibited temporal variation due to water quality changes over time and varied with depth. In the full-scale biofilters, inter-filter variations in the bacterial communities were at least in part due to differences in feed water quality. The abundance of AOB correlated positively with influent ammonia concentration in the pilot-scale biofilters. Filter media type (GAC vs. anthracite) and backwash strategy (chloraminated

vs. non-chloraminated) significantly affected the bacterial community in the pilot-scale biofilters, which provide potential methods of managing the biofilters microbiome. Despite the abovementioned factors, approximately 50% of the variance in the bacterial community was not explained, which could be attributed to other water quality parameters, such as DO, phosphate and nitrate, and upstream treatment processes. Future research should focus on the effects of these factors on the bacterial community in biofilters.

The pilot-scale biofilters effectively removed biomass from the filter influent but only marginally impacted the bacterial community composition in the filter effluent [Chapter 3 and 5]. Although the bacterial communities in the filter effluent were relatively similar to those in the filter influent, there was a clear, albeit minor, impact of the biofilters on the bacterial communities in the filter effluent. Grouping bacteria based on their abundance and occupancy in different samples highlights the organism specific selection in collecting and shedding. 'Influent specific' bacteria were more likely to be captured by the biofilters than bacteria in other groups, whereas 'media specific' and 'media-effluent shared' populations were more likely to be shed from the biofilters. Shedding from the biofilters served as the primary source of AOB in the filter effluent primarily when the AOB concentration in the biofilters was high (> 7 log *amoA*-AOB gene copies/cm^3 bed volume), suggesting that for certain bacteria, the impact of biofilter shedding was related to its abundance in the biofilters. These results point to the need for approaches to manage the microbiome in biofilters to help control problematic bacteria (e.g., opportunistic

pathogens, and corrosion-inducing or odor-causing bacteria) in the finished water and drinking water distribution system.

The microbiomes in full-scale biofilters across North America exhibited a significant distance-decay relationship [Chapter 4]. Distance-decay relationships are normally used to explain the biogeographic patterns of bacterial communities in natural ecosystems [93, 103–105]. According to this relationship, community similarity decreases with geographic distance. This work answered an important ecological question, namely, whether such a relationship can be applied to ecosystems in an engineered system. The distance-decay relationship in the biofilters was significant, but weaker than in natural ecosystems, indicating that both 'nature' (i.e., source water quality and microbiome) and 'nurture' (i.e., engineering decisions) factors influenced the bacterial communities in biofilters.

***Nitrosomonas oligotropha*-like AOB were generally more abundant than AOA in biofilters, and ratio of NOB to AOB generally exceeded the theoretical ratio for conventional two step nitrification [Chapter 5].** *N. oligotropha*-like AOB outnumbered AOA in most full-scale biofilters (12 of 14 biofilters) and in the pilot-scale biofilters over the entire course of the experiment. AOB in the biofilters were comprised primarily of uncultured populations in family *Nitrosomonadaceae* and genus *Nitrosomonas*. For samples with a high abundance of uncultured *Nitrosomonadaceae* (i.e., most full-scale biofilter media samples), the qPCR assay targeting *N. oligotropha* specific *amoA*-AOB gene seriously underestimated the abundance of bacteria currently known as AOB. The

NOB/AOB ratios exceeded the theoretical ratio for conventional two step nitrification (0.5) in most full-scale biofilters (12 of 14 biofilters) and in the pilot-scale biofilters for most of the operation. This was likely attributed to the presence of comammox bacteria (i.e., certain *Nitrospira spp.*, [122, 123]) and the mixotrophic lifestyle of some NOB, which do not rely on nitrification to gain energy for growth. More research is needed to determine the roles of uncultured *Nitrosomonadaceae* and comammox *Nitrospira* in microbial nitrification in water treatment biofilters.

References

1. Benotti MJ, Trenholm R a, Vanderford BJ, Holady JC, Stanford BD, Snyder S a. Pharmaceuticals and endocrine disrupting compounds in U.S. drinking water. *Environ Sci Technol* 2009; **43**: 597–603.

2. Wang C, Shi H, Adams CD, Gamagedara S, Stayton I, Timmons T, et al. Investigation of pharmaceuticals in Missouri natural and drinking water using high performance liquid chromatography-tandem mass spectrometry. *Water Res* 2011; **45**: 1818–1828.

3. Westerhoff P, Yoon Y, Snyder S, Wert E. Fate of endocrine-disruptor, pharmaceutical, and personal care product chemicals during simulated drinking water treatment processes. *Environ Sci Technol* 2005; **39**: 6649–6663.

4. Serrano D, Suárez S, Lema JM, Omil F. Removal of persistent pharmaceutical micropollutants from sewage by addition of PAC in a sequential membrane bioreactor. *Water Res* 2011; **45**: 5323–5333.

5. Snyder SA, Adham S, Redding AM, Cannon FS, DeCarolis J, Oppenheimer J, et al. Role of membranes and activated carbon in the removal of endocrine disruptors and pharmaceuticals. *Desalination* 2007; **202**: 156–181.

6. Kim S, Chu KH, Al-Hamadani Y, Park CM, Jang M, Kim D-H, et al. Removal of contaminants of emerging concern by membranes in water and wastewater: A review. *Chem Eng J* 2018; **335**: 896–914.

7. Salimi M, Esrafili A, Gholami M, Jafari AJ, Kalantary RR, Farzadkia M, et al. Contaminants of emerging concern: a review of new approach in AOP technologies. *Environ Monit Assess* 2017; **189**: 414.

8. Hozalski RM, Goel S, Bouwer EJ. Use of Biofiltration for Removal of Natural Organic Matter to Achieve Biologically Stable Drinking Water. *Water Sci Technol* 1992; **26**: 9–11.

9. Hozalski RM, Goel S, Bouwer EJ. TOC removal in biological filters. *J / Am Water Work Assoc* 1995; **87**: 40–54.

10. Scharf RG, Johnston RW, Semmens MJ, Hozalski RM. Comparison of batch sorption tests, pilot studies, and modeling for estimating GAC bed life. *Water Res* 2010; **44**: 769–780.

11. Ma B, Arnold WA, Hozalski RM. The relative roles of sorption and biodegradation

in the removal of contaminants of emerging concern (CECs) in GAC-sand biofilters. *Water Res* 2018; **146**: 67–76.

12. Collins MR, Eighmy TT, Fenstermacher Jr JM, Spanos SK. Removing natural organic matter by conventional slow sand filtration. *J Am Water Works Assoc* 1992; 80–90.

13. LaPara TM, Wilkinson KH, Strait JM, Hozalski RM, Sadowksy MJ, Hamilton MJ. The bacterial communities of full-scale biologically active, granular activated carbon filters are stable and diverse and potentially contain novel ammonia-oxidizing microorganisms. *Appl Environ Microbiol* 2015; **81**: 6864–6872.

14. Pinto AJ, Xi C, Raskin L. Bacterial Community Structure in the Drinking Water Microbiome Is Governed by Filtration Processes. *Environ Sci Technol* 2012; **46**: 8851–8859.

15. Vignola M, Werner D, Wade MJ, Meynet P, Davenport RJ. Medium shapes the microbial community of water filters with implications for effluent quality. *Water Res* 2018; **129**: 499–508.

16. Bal Krishna KC, Sathasivan A, Chandra Sarker D. Evidence of soluble microbial products accelerating chloramine decay in nitrifying bulk water samples. *Water Res* 2012; **46**: 3977–3988.

17. Sathasivan A, Fisher J, Kastl G. Simple method for quantifying microbiologically assisted chloramine decay in drinking water. *Environ Sci Technol* 2005; **39**: 5407–5413.

18. Melin ES, Ødegaard H. The effect of biofilter loading rate on the removal of organic ozonation by-products. *Water Res* 2000; **34**: 4464–4476.

19. Zhang P, LaPara TM, Goslan EH, Xie Y, Parsons SA, Hozalski RM. Biodegradation of haloacetic acids by bacterial isolates and enrichment cultures from drinking water systems. *Environ Sci Technol* 2009; **43**: 3169–3175.

20. Ho L, Hoefel D, Bock F, Saint CP, Newcombe G. Biodegradation rates of 2-methylisoborneol (MIB) and geosmin through sand filters and in bioreactors. *Chemosphere* 2007; **66**: 2210–2218.

21. Zearley TL, Summers RS. Removal of trace organic micropollutants by drinking water biological filters. *Environ Sci Technol* 2012; **46**: 9412–9419.

22. Kennedy AM, Reinert AM, Knappe DRU, Ferrer I, Summers RS. Full- and pilot-

scale GAC adsorption of organic micropollutants. *Water Res* 2015; **68**: 238–248.

23. Hallé C, Huck PM, Peldszus S. Emerging contaminant removal by biofiltration: temperature , concentration , and EBCT impacts. *J Am Water Works Assoc* 2015; **107**: E364–E379.

24. Zhang S, Gitungo SW, Axe L, Raczko RF, Dyksen JE. Biologically active filters - An advanced water treatment process for contaminants of emerging concern. *Water Res* 2017; **114**: 31–41.

25. Tunkel J, Howard PH, Boethling RS, Stiteler W, Loonen H. Predicting ready biodegradability in the Japanese Ministry of International Trade and Industry test. *Environ Toxicol Chem* 2000; **19**: 2478–2485.

26. Graham MR, Summers RS, Simpson MR, MacLeod BW. Modeling equilibrium adsorption of 2-methylisoborneol and geosmin in natural waters. *Water Res* 2000; **34**: 2291–2300.

27. Vanderford BJ, Snyder SA. Analysis of pharmaceuticals in water by isotope dilution liquid chromatography/tandem mass spectrometry. *Environ Sci Technol* 2006; **40**: 7312–7320.

28. Trenholm RA, Vanderford BJ, Holady JC, Rexing DJ, Snyder SA. Broad range analysis of endocrine disruptors and pharmaceuticals using gas chromatography and liquid chromatography tandem mass spectrometry. *Chemosphere* 2006; **65**: 1990–1998.

29. Gao J, Ellis LBM, Wackett LP. The University of Minnesota Biocatalysis/Biodegradation Database: Improving public access. *Nucleic Acids Res* 2009; **38**: 488–491.

30. Helbling DE, Hollender J, Kohler H-PE, Singer H, Fenner K. SI-High-throughput identification of microbial transformation products of organic micropollutants. *Environ Sci Technol* 2010; **44**: 6621–7.

31. Ridder DJ De, Verliefde ARD, Heijman SGJ, Verberk JQJC, Rietveld LC, Aa LTJ Van Der, et al. Influence of natural organic matter on equilibrium adsorption of neutral and charged pharmaceuticals onto activated carbon. 2011; 416–423.

32. Nam S-W, Choi D-J, Kim S-K, Her N, Zoh K-D. Adsorption characteristics of selected hydrophilic and hydrophobic micropollutants in water using activated carbon. *J Hazard Mater* 2014; **270**: 144–52.

33. Summers RS, Shimabuku K, Zearley TL. A Review of Biologically-Based Drinking Water Treatment Processes for Organic Micropollutant Removal. *Prog Slow Sand Altern Biofiltration Process Furth Dev Appl* 2014; 17–24.

34. Reungoat J, Escher BI, Macova M, Keller J. Biofiltration of wastewater treatment plant effluent: Effective removal of pharmaceuticals and personal care products and reduction of toxicity. *Water Res* 2011; **45**: 2751–2762.

35. Carpenter CMG, Helbling DE. Removal of micropollutants in biofilters: Hydrodynamic effects on biofilm assembly and functioning. *Water Res* 2017; **120**: 211–221.

36. Radjenović J, Pérez S, Petrović M, Barceló D. Identification and structural characterization of biodegradation products of atenolol and glibenclamide by liquid chromatography coupled to hybrid quadrupole time-of-flight and quadrupole ion trap mass spectrometry. *J Chromatogr A* 2008; **1210**: 142–153.

37. Liu D, Maguire RJ, Pacepavicius GJ. Microbial Transformation of Metolachlor. *Envirionmental Toxicol Water Qual An Int J* 1995; **10**: 249–258.

38. Aga DS, Thurman EM. Formation and transport of the sulfonic acid metabolites of alachlor and metolachlor in soil. *Environ Sci Technol* 2001; **35**: 2455–2460.

39. Velázquez YF, Nacheva PM. Biodegradability of fluoxetine, mefenamic acid, and metoprolol using different microbial consortiums. *Environ Sci Pollut Res* 2017; **24**: 6779–6793.

40. Ternes TA, Meisenheimer M, McDowell D, Sacher F, Brauch HJ, Haist-Gulde B, et al. Removal of pharmaceuticals during drinking water treatment. *Environ Sci Technol* 2002; **36**: 3855–3863.

41. Corwin CJ, Summers RS. Adsorption and desorption of trace organic contaminants from granular activated carbon adsorbers after intermittent loading and throughout backwash cycles. *Water Res* 2011; **45**: 417–426.

42. Magnuson ML, Speth TF. Quantitative structure--property relationships for enhancing predictions of synthetic organic chemical removal from drinking water by granular activated carbon. *Env Sci Technol* 2005; **39**: 7706–7711.

43. Quinlivan PA, Li L, Knappe DRU. Effects of activated carbon characteristics on the simultaneous adsorption of aqueous organic micropollutants and natural organic matter. *Water Res* 2005; **39**: 1663–1673.

44. Newcombe G, Hayes R, Drikas M. Granular activated carbon: Importance of surface properties in the adsorption of naturally occurring organics. *Colloids Surfaces A Physicochem Eng Asp* 1993; **78**: 65–71.

45. Bjelopavlic M, Newcombe G, Hayes R. Adsorption of NOM onto Activated Carbon: Effect of Surface Charge, Ionic Strength, and Pore Volume Distribution. *J Colloid Interface Sci* 1999; **210**: 271–280.

46. Fairey JL, Speitel GE, Katz LE. Impact of natural organic matter on monochloramine reduction by granular activated carbon: The role of porosity and electrostatic surface properties. *Environ Sci Technol* 2006; **40**: 4268–4273.

47. Anumol T, Sgroi M, Park M, Roccaro P, Snyder SA. Predicting trace organic compound breakthrough in granular activated carbon using fluorescence and UV absorbance as surrogates. *Water Res* 2015; **76**: 76–87.

48. Kennedy AM, Summers RS. Effect of DOM size on organic micropollutant adsorption by GAC. *Environ Sci Technol* 2015; **49**: 6617–6624.

49. Emelko MB, Huck PM, Coffey BM, Smith EF. Effects of media, backwash, and temperature on full-scale biological filtration. *J / Am Water Work Assoc* 2006; **98**: 61–73.

50. Corwin CJ, Summers RS. Scaling trace organic contaminant adsorption capacity by granular activated carbon. *Environ Sci Technol* 2010; **44**: 5403–8.

51. Niquette P, Servais P, Savoir R. Impacts of pipe materials on densities of fixed bacterial biomass in a drinking water distribution system. *Water Res* 2000; **34**: 1952–1956.

52. Vieno NM, Härkki H, Tuhkanen T, Kronberg L. Occurrence of pharmaceuticals in river water and their elimination in a pilot-scale drinking water treatment plant. *Environ Sci Technol* 2007; **41**: 5077–5084.

53. Khunjar WO, MacKintosh SA, Skotnicka-Pitak J, Baik S, Aga DS, Love NG. Elucidating the relative roles of ammonia oxidizing and heterotrophic bacteria during the biotransformation of 17α-ethinylestradiol and trimethoprim. *Environ Sci Technol* 2011; **45**: 3605–3612.

54. Torsvik V, Øvreås L, Thingstad TF. Prokaryotic Diversity — Magnitude , Dynamics , and Controlling Factors. 2002; **296**: 1064–1067.

55. Nemergut DR, Schmidt SK, Fukami T, Neill SPO, Bilinski TM, Stanish LF, et al.

Patterns and Processes of Microbial Community Assembly. *Microbiol Mol Biol Rev* 2013; **77**: 342–356.

56. Hubbell SP. Neutral theory in community ecology and the hypothesis of functional equivalence. *Funct Ecol* 2005; **19**: 166–172.

57. Martiny JBH, Bohannan BJM, Brown JH, Kane M, Krumins JA, Kuske CR, et al. Microbial biogeography : putting microorganisms on the map. *Nat Rev Microbiol* 2006; **4**: 102–112.

58. Grenni P, Patrolecco L, Ademollo N, Tolomei A, Barra Caracciolo A. Degradation of Gemfibrozil and Naproxen in a river water ecosystem. *Microchem J* 2013; **107**: 158–164.

59. Wunder DB, Tan DT, LaPara TM, Hozalski RM. The effects of antibiotic cocktails at environmentally relevant concentrations on the community composition and acetate biodegradation kinetics of bacterial biofilms. *Chemosphere* 2013; **90**: 2261–2266.

60. Aggarwal S, Gomez-Smith CK, Jeon Y, Lapara TM, Waak MB, Hozalski RM. Effects of Chloramine and Coupon Material on Biofilm Abundance and Community Composition in Bench-Scale Simulated Water Distribution Systems and Comparison with Full-Scale Water Mains. *Environ Sci Technol* 2018; **52**: 13077–13088.

61. Gomez-Smith CK, Lapara TM, Hozalski RM. Sulfate reducing bacteria and mycobacteria dominate the biofilm communities in a chloraminated drinking water distribution system. *Environ Sci Technol* 2015; **49**: 8432–8440.

62. Muyzer G, de Waal E, Uitterlinden AG. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain. *Appl Environ Microbiol* 1993; **59**: 695–700.

63. Muyzer G, De Waal EC, Uitterlinden AG. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* 1993; **59**: 695–700.

64. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* 2016; **13**: 581–583.

65. Callahan BJ, McMurdie PJ, Holmes SP. Exact sequence variants should replace

operational taxonomic units in marker-gene data analysis. *ISME J* 2017; **11**: 2639–2643.

66. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2012; **41**: 590–596.

67. Price M, Dehal P, Arkin A. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS One* 2010; **5**: e9490.

68. Ling F, Hwang C, LeChevallier MW, Andersen GL, Liu WT. Core-satellite populations and seasonality of water meter biofilms in a metropolitan drinking water distribution system. *ISME J* 2016; **10**: 582–595.

69. Pinto AJ, Schroeder JL, Lunn M, Sloan WT, Raskin L. Spatial-Temporal Survey and Occupancy-Abundance Modeling To Predict Bacterial Community Dynamics in the Drinking Water Microbiome. *MBio* 2014; **5**: 1–13.

70. Jean Dunn O. Multiple Comparisons Among Means. *J Am Stat Assoc* 1961; **56**: 52–64.

71. Benjimini Y, Hochberg Y. Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. *J R Stat Soc B* 1995; **57**: 289–300.

72. Hahn MW, Kasalický V, Jezbera J, Brandt U, Jezberová J, Šimek K. Limnohabitans curvus gen. nov., sp. nov., a planktonic bacterium isolated from a freshwater lake. *Int J Syst Evol Microbiol* 2010; **60**: 1358–1365.

73. Willems A, Busse J, Goor M, Pot B, Falsen E, Jantzen E, et al. Hydrogenophaga, a New Genus of Hydrogen-Oxidizing Bacteria That Includes Hydrogenophaga flava comb. nov. (Formerly Pseudomonas flava), Hydrogenophaga palleronii (Formerly Pseudomonas palleronii), Hydrogenophaga pseudoflava (Formerly Pseudomonas pseudoflav. *Int J Syst Bacteriol* 1989; 319–333.

74. Wang NF, Zhang T, Zhang F, Wang ET, He JF, Ding H, et al. Diversity and structure of soil bacterial communities in the Fildes Region (maritime Antarctica) as revealed by 454 pyrosequencing. *Front Microbiol* 2015; **6**: 1–11.

75. Zwart G, Crump BC, Kamst-van Agterveld MP, Hagen F, Han SK. Typical freshwater bacteria: An analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. *Aquat Microb Ecol* 2002; **28**: 141–155.

76. Wexler HM. Bacteroides: The good, the bad, and the nitty-gritty. *Clin Microbiol*

Rev 2007; **20**: 593–621.

77. Reynolds K, Kristina M, Gerba C. Risk of Waterborne Illness Via Drinking Water in the United States. *Rev Environ Contam Toxicol* 2008; **236**: 117–158.

78. Proctor CR, Hammes F. Drinking water microbiology-from measurement to management. *Curr Opin Biotechnol* 2015; **33**: 87–94.

79. Rittmann B. The Effect of Shear Stress on Biofilm Loss Rate. *Biotechnol Bioeng* 1982.

80. Harald H, Hempel D. Growth and Decay in an Auto-/Heterotrophic Biofilm. *Water Res* 1997; **31**: 2243–2252.

81. Irgens RL, Gosink JJ, Staley JT. Polaromonas vacuolata gen. nov., sp. nov., a Psychrophilic, Marine, Gas Vacuolate Bacterium from Antarctica. *Int J Syst Bacteriol* 1996; **46**: 822–826.

82. Lee EM, Jeon CO, Choi I, Chang KS, Kim CJ. Silanimonas lenta gen. nov., sp. nov., a slightly thermophilic and alkaliphilic gammaproteobacterium isolated from a hot spring. *Int J Syst Evol Microbiol* 2005; **55**: 385–389.

83. Kasalický V, Jezbera J, Šimek K, Hahn MW. Limnohabitans planktonicus sp. nov. and Limnohabitans parvus sp. nov., planktonic betaproteobacteria isolated from a freshwater reservoir, and emended description of the genus Limnohabitans. *Int J Syst Evol Microbiol* 2010; **60**: 2710–2714.

84. Helbling DE, Johnson DR, Honti M, Fenner K. Micropollutant biotransformation kinetics associate with WWTP process parameters and microbial community characteristics. *Environ Sci Technol* 2012; **46**: 10579–10588.

85. Hartgers EM, Aalderink GH, Van Den Brink PJ, Gylstra R, Wiegman JWF, Brock TCM. Ecotoxicological threshold levels of a mixture of herbicides (atrazine, diuron and metolachlor) in freshwater microcosms. *Aquat Ecol* 1998; **32**: 135–152.

86. Johansson CH, Janmar L, Backhaus T. Toxicity of ciprofloxacin and sulfamethoxazole to marine periphytic algae and bacteria. *Aquat Toxicol* 2014; **156**: 248–258.

87. Jos A, Repetto G, Rios JC, Hazen MJ, Molero ML, Del Peso A, et al. Ecotoxicological evaluation of carbamazepine using six different model systems with eighteen endpoints. *Toxicol Vitr* 2003; **17**: 525–532.

88. De Castro-Català N, Muñoz I, Riera JL, Ford AT. Evidence of low dose effects of the antidepressant fluoxetine and the fungicide prochloraz on the behavior of the keystone freshwater invertebrate Gammarus pulex. *Environ Pollut* 2017; **231**: 406–414.

89. De Vera GA, Gerrity D, Stoker M, Frehner W, Wert EC. Impact of upstream chlorination on filter performance and microbial community structure of GAC and anthracite biofilters. *Environ Sci Water Res Technol* 2018; **4**: 1133–1144.

90. Liao X, Chen C, Zhang J, Dai Y, Zhang X, Xie S. Operational performance, biomass and microbial community structure: Impacts of backwashing on drinking water biofilter. *Environ Sci Pollut Res* 2014; **22**: 546–554.

91. Bauer RC, Snoeyink VL. Reactions of chloramines with active carbon. *J Water Pollut Control Fed* 1973; **45**: 2290–2301.

92. Velten S, Boller M, Koster O, Helbing J, Weilenmann HU, Hammes F. Development of biomass in a drinking water granular active carbon (GAC) filter. *Water Res* 2011; **45**: 6347–6354.

93. Fierer N, Jackson R. The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci U S A* 2006; **103**: 626–631.

94. Lu XL, Wang N, Wang HM, Deng YM, Ma T, Wu MXJ, et al. Molecular Characterization of the Total Bacteria and Dissimilatory Arsenate-Reducing Bacteria in Core Sediments of the Jianghan Plain, Central China. *Geomicrobiol J* 2017; **34**: 467–479.

95. Loreau M, Naeem S, Inchausti P, Bengtsson J, Grime JP, Hector A, et al. Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science (80-)* 2001; **294**: 804–808.

96. Allison S, Martiny J. Resistance, resilience, and redundancy in microbial communities. *Proc Natl Acad Sci U S A* 2008; **105**: 11512–11519.

97. Hosen JD, Febria CM, Crump BC, Palmer MA. Watershed urbanization linked to differences in stream bacterial community composition. *Front Microbiol* 2017; **8**: 1–17.

98. Douterelo I, Sharpe R., Boxall JB. Influence of hydraulic regimes on bacterial community structure and composition in an experimental drinking water distribution system. *Water Res* 2013; **47**: 503–516.

99. Donlan RM. Biofilms: microbial life on surfaces. Emerg Infect Diseases. *Emerg Infect Dis* 2002; **8**: 881–890.

100. Whitchurch CB, Tolker-Nielsen T, Ragas PC, Mattick JS. Extracellular DNA required for bacterial biofilm formation. *Science (80-)* 2002; **295**: 1487.

101. O'Toole G, Kaplan HB, Kolter R. Biofilm Formation as Microbial Development. *Annu Rev Microbiol* 2000; **54**: 49–79.

102. Darby JL, Lawler DF. Ripening in Depth Filtration: Effect of Particle Size on Removal and Head Loss. *Environ Sci Technol* 1990; **24**: 1069–1079.

103. Astorga A, Oksanen J, Luoto M, Soininen J, Virtanen R, Muotka T. Distance decay of similarity in freshwater communities: Do macro- and microorganisms follow the same rules? *Glob Ecol Biogeogr* 2012; **21**: 365–375.

104. Green J, Holmes A, Westoby M, Oliver I, Briscoe D, Dangerfield M, et al. Spatial scaling of microbial eukaryote diversity. *Nature* 2004; **432**: 747–750.

105. Martiny JBH, Eisen JA, Penn K, Allison SD, Horner-devine MC. Drivers of bacterial β -diversity depend on spatial scale. *Proc Natl Acad Sci U S A* 2011; **108**: 7850–7854.

106. Nekola J, White P. The distance decay of similarity in biogeography and ecology. *J Biogeogr* 1999; **26**: 867–878.

107. Wang X, Wen X, Deng Y, Xia Y, Yang Y, Zhou J. Distance-decay relationship for biological wastewater treatment plants. *Appl Environ Microbiol* 2016; **82**: 4860–4866.

108. Zhang B, Xia Y, Wen X, Wang X, Yang Y, Zhou J, et al. The composition and spatial patterns of bacterial virulence factors and antibiotic resistance genes in 19 wastewater treatment plants. *PLoS One* 2016; **11**: 1–14.

109. Horner-Devine MCC, Lage M, Hughes JBB, Bohannan BJMJM. Is there a species-area relationship for bacteria? *Nature* 2004; **432**: 750–753.

110. Lichstein JW. Multiple regression on distance matrices : a multivariate spatial analysis tool. *Plant Ecol* 2007; **188**: 117–131.

111. Ohkouchi Y, Ly BT, Ishikawa S, Kawano Y, Itoh S. Determination of an acceptable assimilable organic carbon (AOC) level for biological stability in water distribution systems with minimized chlorine residual. *Environ Monit Assess* 2013; **185**: 1427–

1436.

112. Waak MB, Hozalski RM, Hallé C, Lapara TM. Comparison of the microbiomes of two drinking water distribution systems — with and without residual chloramine disinfection. *Microbiome* 2019.

113. Hanson CA, Fuhrman JA, Horner-Devine MC, Martiny JBH. Beyond biogeographic patterns: Processes shaping the microbial landscape. *Nat Rev Microbiol* 2012; **10**: 497–506.

114. Green J, Bohannan BJM. Spatial scaling of microbial biodiversity. *Trends Ecol Evol* 2006; **21**: 501–507.

115. Ameet J. Pinto CX and LR. Bacterial Community Structure in the Drinking Water Microbiome Is Governed by Filtration Processes. *Environ Sci Technol* 2012; **46**: 8851−8859.

116. Ma B, Lapara TM, Hozalski RM. Bacterial communities in drinking water biofilters exhibit temporal variation, vary with media type, backwash strategy, and bed depth, and have a minor impact on the effluent water microbiome. 2019.

117. Prosser JI, Head IM, Stein LY. The Family Nitrosomonadaceae. *The Prokaryotes*. 2014.

118. Purkhold U, Pommerening-ro A, Juretschko S, Schmid M, Hans-Peter K, Wagner M. Phylogeny of All Recognized Species of Ammonia Oxidizers Based on Comparative 16S rRNA and amoA Sequence Analysis: Implications for Molecular Diversity Surveys. *Appl Envir Microbiol* 2000; **66**: 5368–5382.

119. Hatzenpichler R. Diversity, Physiology, and Niche Differentiation of Ammonia-Oxidizing Archaea. *Appl Environ Microbiol* 2012; **78**: 7501–7510.

120. Abeliovich A. The Nitrite-Oxidizing Bacteria. *Prokaryotes*. 2006. pp 861–872.

121. Daims H, Lücker S, Wagner M. A New Perspective on Microbes Formerly Known as Nitrite-Oxidizing Bacteria. *Trends Microbiol* 2016; **24**: 699–712.

122. Daims H, Lebedeva E V, Pjevac P, Han P, Herbold C, Jehmlich N, et al. Complete nitrification by Nitrospira bacteria. *Nature* 2016; **528**: 504–509.

123. Kessel MAHJ Van, Speth DR, Albertsen M, Nielsen PH, Huub JM. Complete nitrification by a single microorganism. *Nature* 2016; **528**: 555–559.

124. Ren YX, Nakano K, Nomura M, Chiba N, Nishimura O. Effects of bacterial activity on estrogen removal in nitrifying activated sludge. *Water Res* 2007; **41**: 3089–3096.

125. Wahman DG, Kirisits MJ, Katz LE, Speitel GE. Ammonia-oxidizing bacteria in biofilters removing trihalomethanes are related to Nitrosomonas oligotropha. *Appl Environ Microbiol* 2011; **77**: 2537–2540.

126. Bartram AK, Lynch MDJ, Stearns JC, Moreno-Hagelsieb G, Neufeld JD. Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end Illumina reads. *Appl Environ Microbiol* 2011; **77**: 3846–3852.

127. Harms G, Layton AC, Dionisi HM, Gregory IR, Garrett VM, Hawkins SA, et al. Real-time PCR quantification of nitrifying bacteria in a municipal wastewater treatment plant. *Environ Sci Technol* 2003; **37**: 343–351.

128. Meinhardt KA, Bertagnolli A, Pannu MW, Strand SE, Brown SL, Stahl DA. Evaluation of revised polymerase chain reaction primers for more inclusive quantification of ammonia-oxidizing archaea and bacteria. *Environ Microbiol Rep* 2015; **7**: 354–363.

129. Norton JM, Alzerreca JJ, Suwa Y, Klotz MG. Diversity of ammonia monooxygenase operon in autotrophic ammonia-oxidizing bacteria. 2002; 139–149.

130. Layton AC, Dionisi H, Kuo H, Robinson KG, Garrett VM, Meyers A, et al. Emergence of Competitive Dominant Ammonia-Oxidizing Bacterial Populations in a Full-Scale Industrial Wastewater Treatment Plant. 2005; **71**: 1105–1108.

131. Větrovský T, Baldrian P. The Variability of the 16S rRNA Gene in Bacterial Genomes and Its Consequences for Bacterial Community Analyses. *PLoS One* 2013; **8**: 1–10.

132. Hagopian DS, Riley JG. A closer look at the bacteriology of nitrification. *Aquac Eng* 1998; **18**: 223–244.

133. Winkler MKH, Bassin JP, Kleerebezem R, Sorokin DY, Van Loosdrecht MCM. Unravelling the reasons for disproportion in the ratio of AOB and NOB in aerobic granular sludge. *Appl Microbiol Biotechnol* 2012; **94**: 1657–1666.

134. Niu J, Kasuga I, Kurisu F, Furumai H, Shigeeda T. Evaluation of autotrophic growth of ammonia-oxidizers associated with granular activated carbon used for drinking water purification byDNA-stable isotope probing. *Water Res* 2013; **47**: 7053–7065.

135. Van Der Wielen PWJJ, Voost S, Van Der Kooij D. Ammonia-oxidizing bacteria and archaea in groundwater treatment and drinking water distribution systems. *Appl Environ Microbiol* 2009; **75**: 4687–4695.

136. Kasuga I, Nakagaki H, Kurisu F, Furumai H. Predominance of ammonia-oxidizing archaea on granular activated carbon used in a full-scale advanced drinking water treatment plant. *Water Res* 2010; **44**: 5039–5049.

137. de Vet WWJM, Kleerebezem R, van der Wielen PWJJ, Rietveld LC, van Loosdrecht MCM. Assessment of nitrification in groundwater filters for drinking water production by qPCR and activity measurement. *Water Res* 2011; **45**: 4008–4018.

138. Martens-Habbena W, Berube PM, Urakawa H, De La Torre JR, Stahl DA. Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria. *Nature* 2009; **461**: 976–979.

139. Pintar KDM, Slawson RM. Effect of temperature and disinfection strategies on ammonia-oxidizing bacteria in a bench-scale drinking water distribution system. *Water Res* 2003; **37**: 1805–1817.

140. Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, et al. PubChem substance and compound databases. *Nucleic Acids Res* 2016; **44**: D1202–D1213.

141. Eren AM, Faust K, Unterseher M, Lindahl B, Bálint M, Fuhrman JA, et al. Millions of reads, thousands of taxa: microbial community structure and associations analyzed via marker genes. *FEMS Microbiol Rev* 2016; **40**: 686–700.

142. Pollock J, Laura G, Trong W, Mick W. The Madness of Microbiome : Attempting To Find Consensus. *Appl Environ Microbiol* 2018; **84**: 1–12.

143. Gieseke A, Purkhold U, Wagner M, Amann R, Schramm A. Community Structure and Activity Dynamics of Nitrifying Bacteria in a Phosphate-Removing Biofilm. *Appl Envir Microbiol* 2001; **67**: 1351–1362.

144. Dionisi HM, Layton AC, Harms G, Gregory IR, Robinson KG, Sayler GS, et al. Quantification of Nitrosomonas oligotropha -Like Ammonia-Oxidizing Bacteria and Nitrospira spp . from Full-Scale Wastewater Treatment Plants by Competitive PCR Quantification of Nitrosomonas oligotropha-Like Ammonia- Oxidizing Bacteria and Nitrospira spp. *Appl Environ Microbiol* 2002; **68**: 245–253.

145. Gruber-Dorninger C, Pester M, Kitzinger K, Savio DF, Loy A, Rattei T, et al. Functionally relevant diversity of closely related Nitrospira in activated sludge. *ISME J* 2015; **9**: 643–655.

146. Koch H, Lucker S, Albertsen M, Kitzinger K, Herbold C, Spieck E, et al. Expanded metabolic versatility of ubiquitous nitrite-oxidizing bacteria from the genus Nitrospira. *Proc Natl Acad Sci* 2015; **112**: 11371–11376.

Appendix A

Photos and schematic diagrams of the pilot-scale biofiltration system at MWTDS

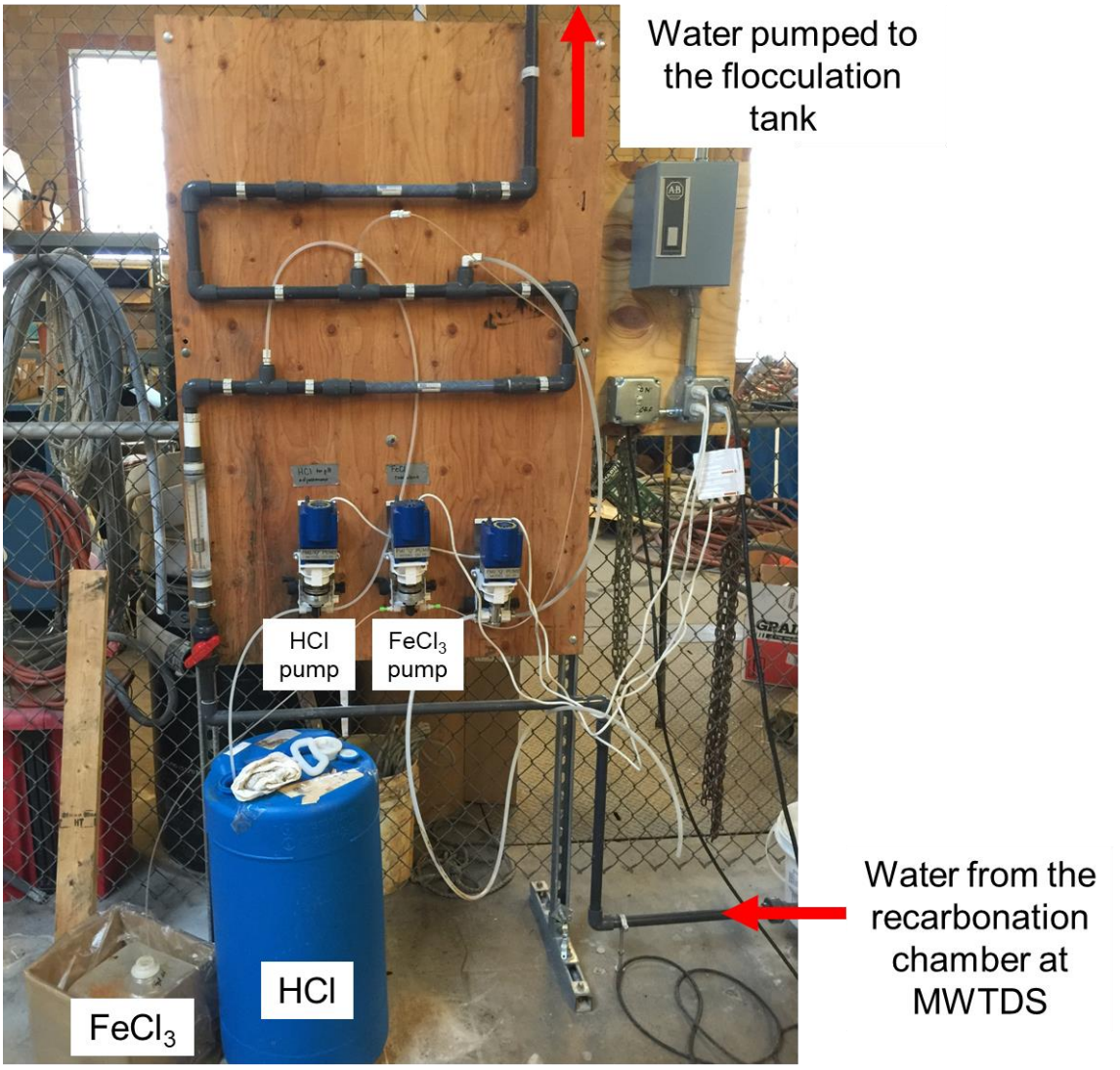

Figure A- 1. Photo of the pumps for HCl and $FeCl_3$ dosing in the pilot-scale biofiltration system at MWTDS.

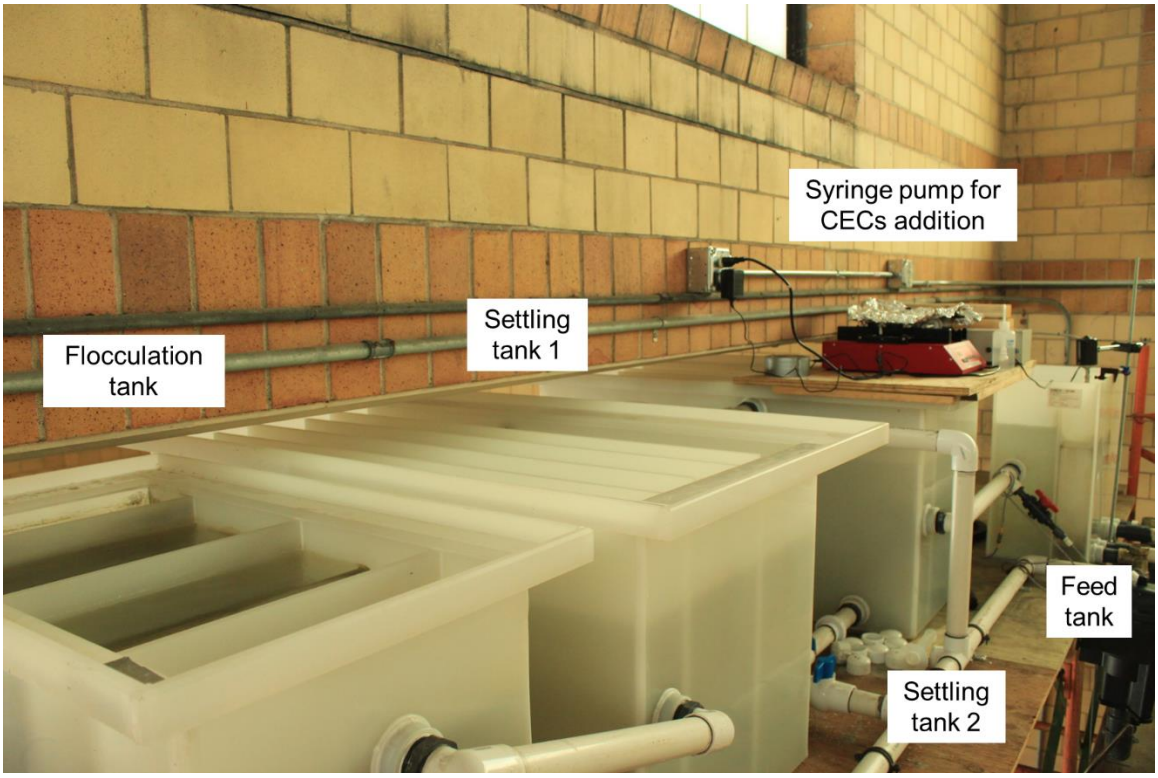

Figure A- 2. Photo of the flocculation tank, two settling tanks, feed tank and syringe pump for CECs addition in the pilot-scale biofiltration system at MWTDS.

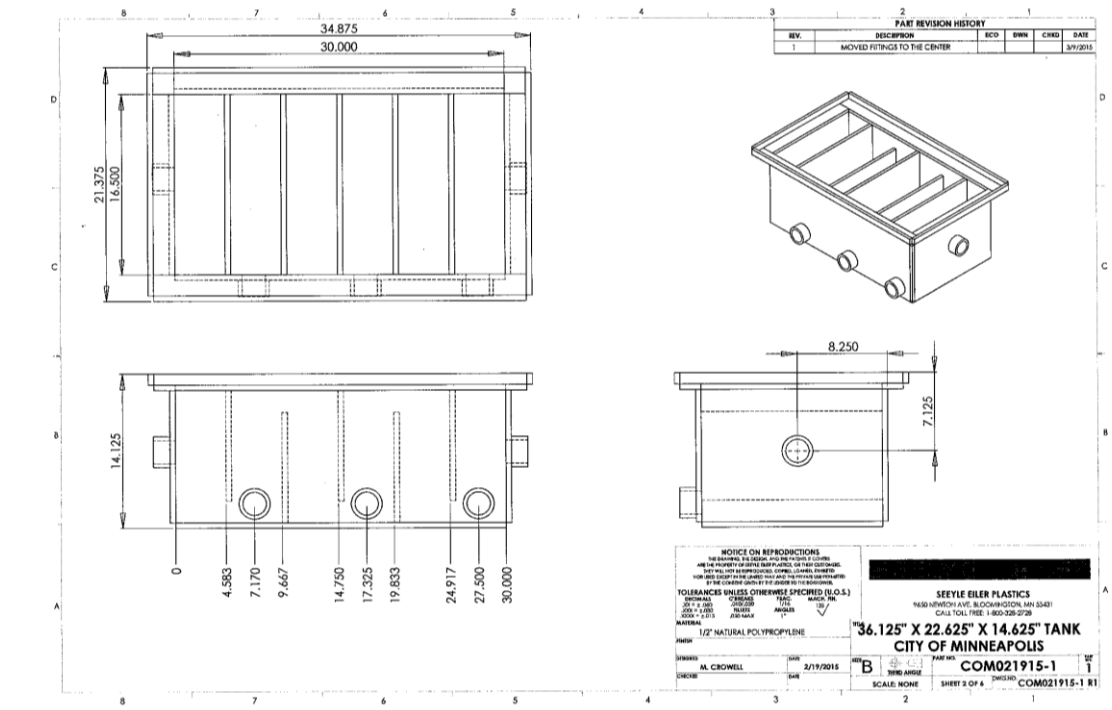

Figure A- 3. Schematic diagram of the flocculation tank in the pilot-scale biofiltration system at MWTDS.

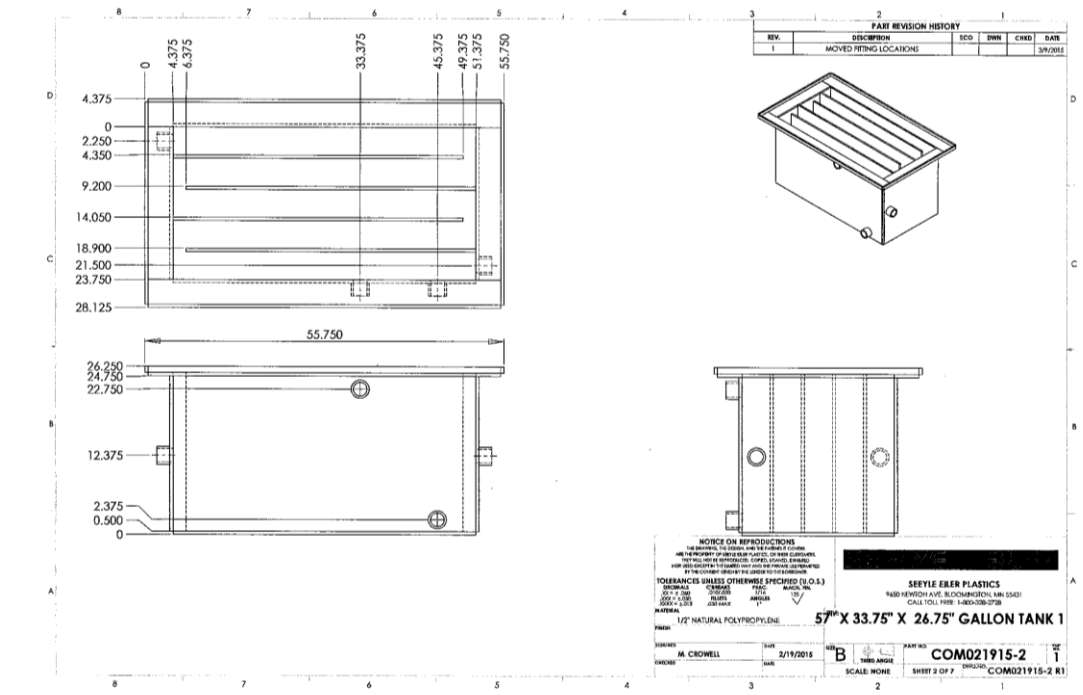

Figure A- 4. Schematic diagram of the settling tank in the pilot-scale biofiltration system at MWTDS.

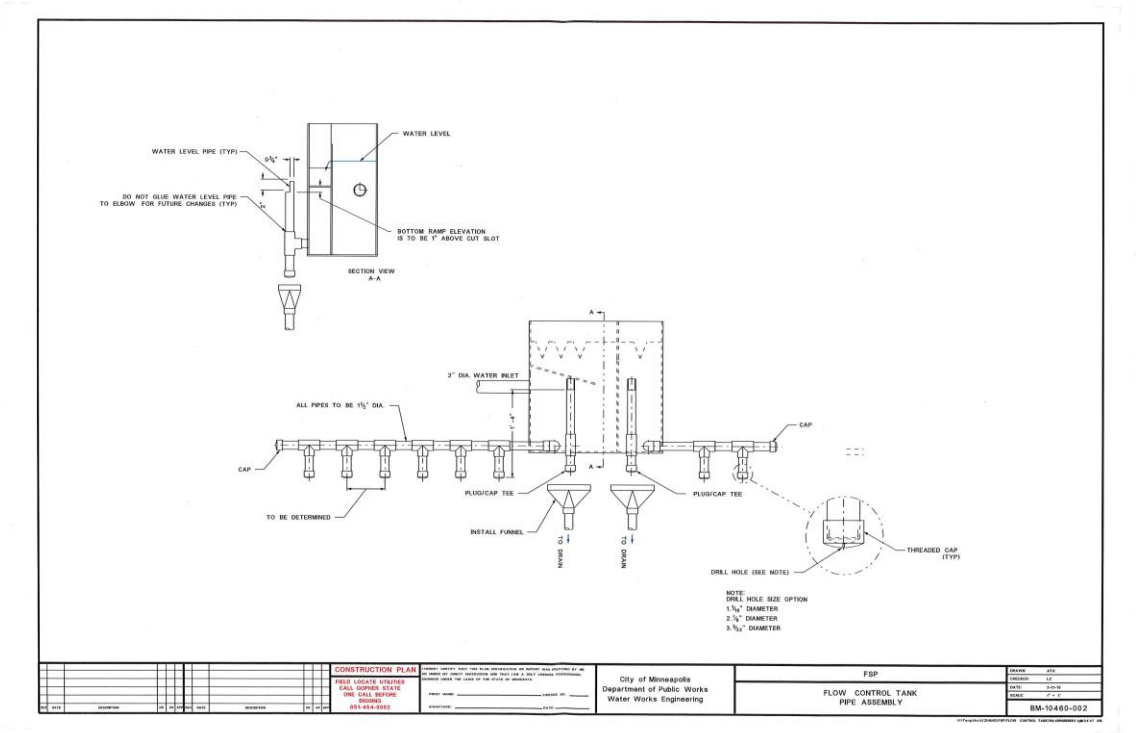

Figure A- 5. Schematic diagram of the feed tank and the filter influent inlet in the pilot-scale biofiltration system at MWTDS.

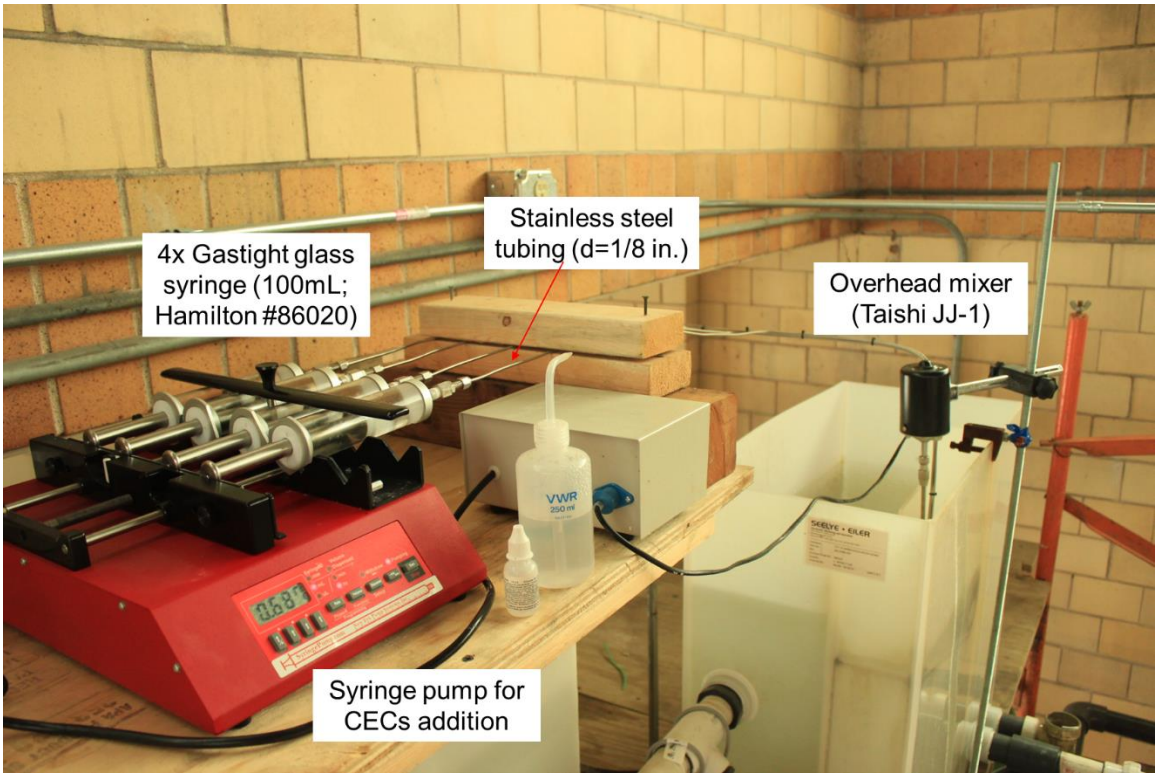

Figure A- 6. Photo of the CECs dosing and mix systems in the pilot-scale biofiltration system at MWTDS.

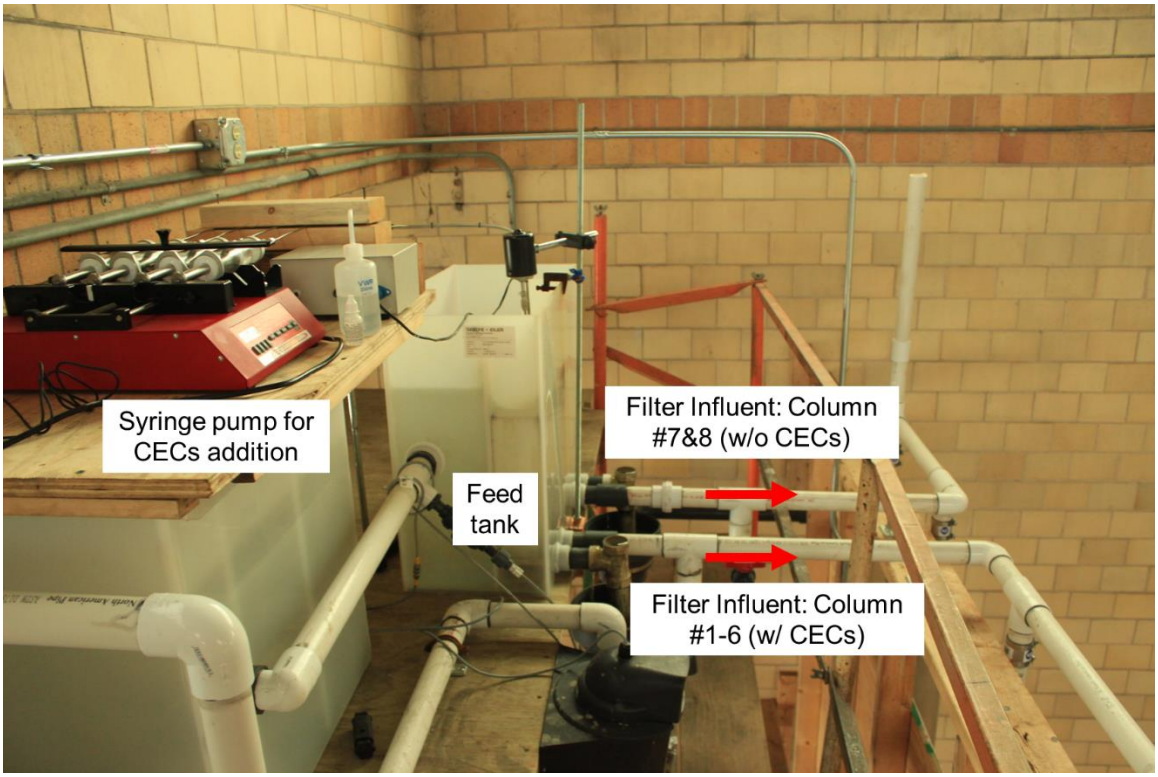

Figure A- 7. Photo of the feed and filter influent inlets in the pilot-scale biofiltration system at MWTDS.

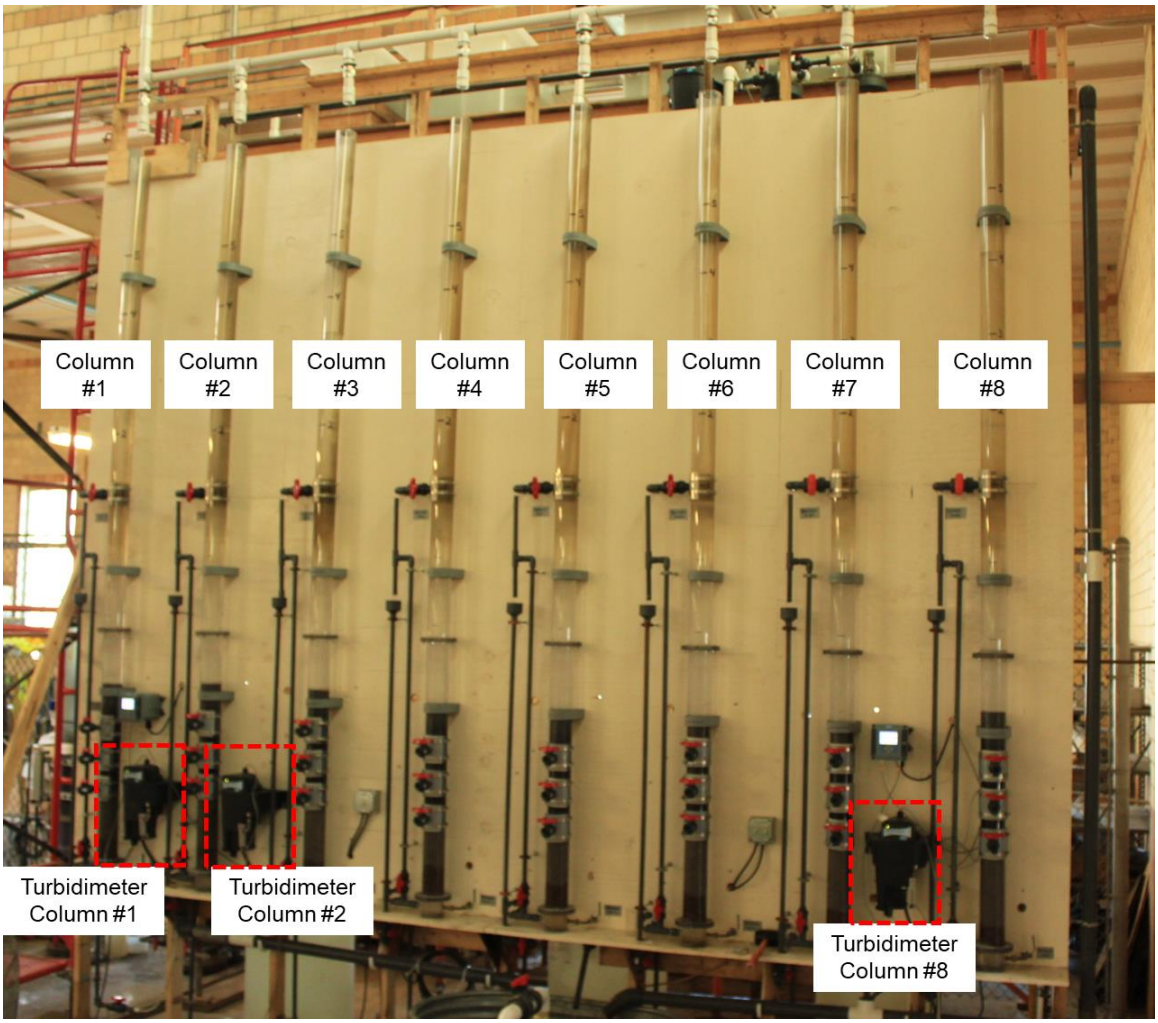

Figure A- 8. Photo of the biofilter columns and turbidimeters for filter effluent turbidity measurement in the pilot-scale biofiltration system at MWTDS.

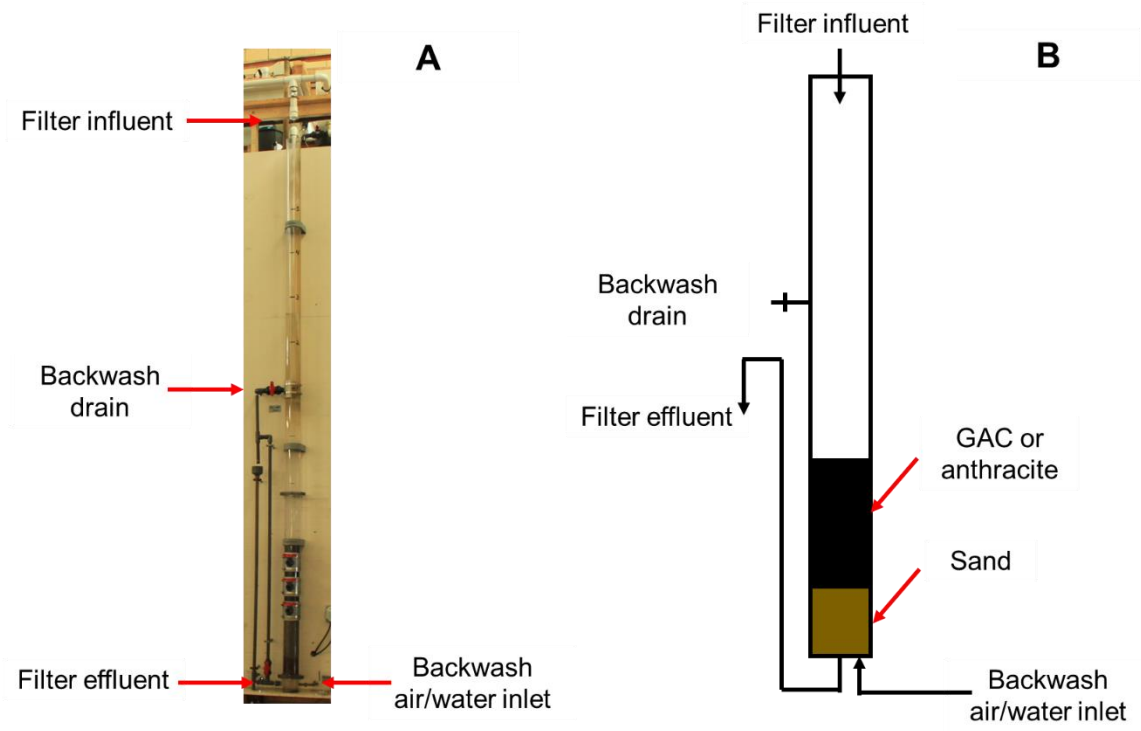

Figure A- 9. Photo (**A**) and schematic diagram (**B**) of the biofilter column in the pilot-scale biofiltration system at MWTDS.

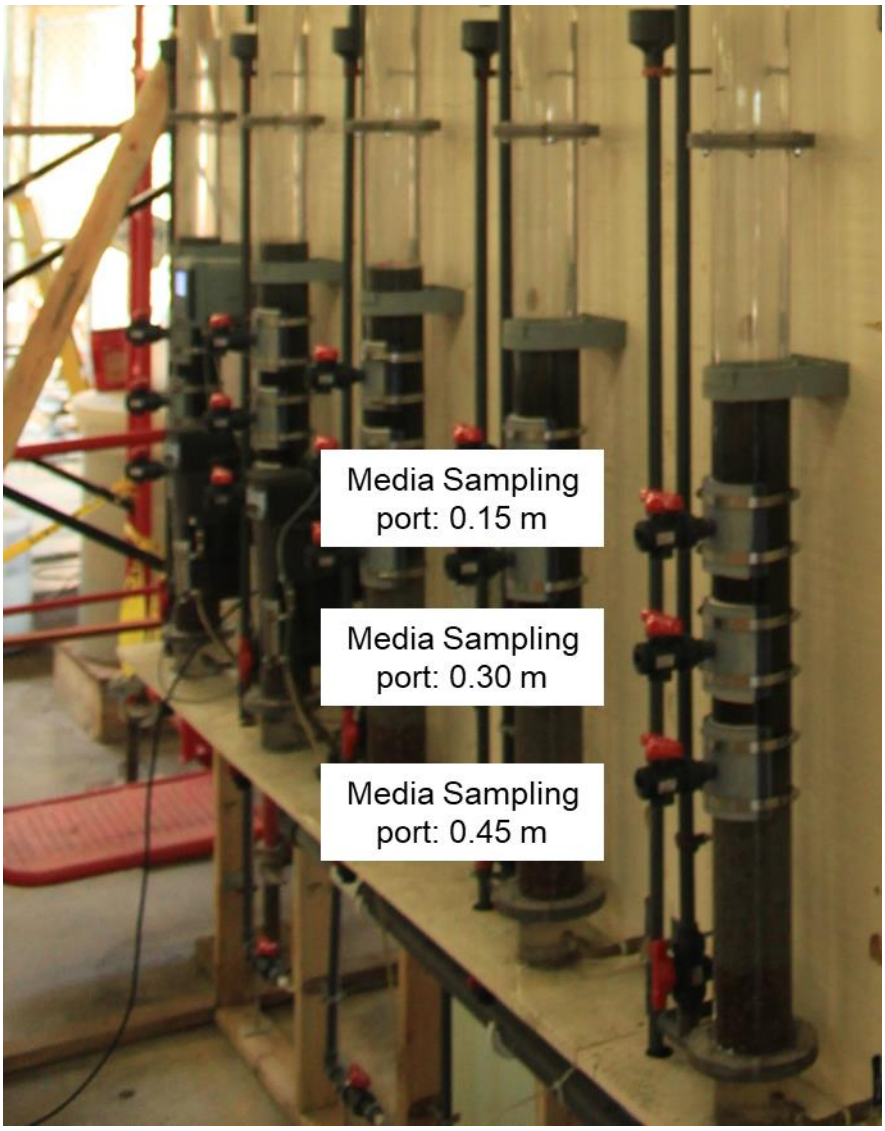

Figure A- 10. Photo of the sampling ports for GAC or anthracite media sample collection from different bed depth (0.15m, 0.30m and 0.45m) in the pilot-scale biofiltration system at MWTDS.

Appendix B

Supporting Information for Chapter 2

B-1. Liquid chromatography tandem mass spectrometry (LC-MS/MS)

This method was adapted from Vanderford and Snyder (2006). A 0.5 mm × 150 mm ZORBAX SB-C18 column with 5 μm particle size (Agilent Technologies) was used to separate the selected CECs. Two different gradient programs were used in LC: one for for ESI positive analytes (1) and another for ESI negative analytes (2) as described below. The gradient was comprised of a blend of 5 mM ammonium acetate in ultrapure water (A) and 100% methanol (B) fed at a total flow rate of 20 μL/min.

(1) For atenolol, atrazine, carbamazepine, fluoxetine, metolachlor, sulfamethoxazole and tris (2-chlorethly) phosphate (ESI positive): 10% B held for 5 min, then gradually increased to 98% over 10 min and held at 98% for 10 min.

(2) For gemfibrozil (ESI negative): started at 50% B, gradually increased to 98% over 2 min and held at 98% for 7 min.

The detailed MS/MS parameters for all selected CECs are summarized in Table S4. All compounds were monitored for the entire length of the run. The limit of detection (LOD) and limit of quantification (LOQ) were calculated by multiplying the standard deviation of analyte concentrations in triplicate method blanks by 3 and 10, respectively. These values are also listed in Table A-4.

Calibration standards for LC-MS/MS were prepared in 0.1 mol/L carbonate buffered ultrapure water (pH 8.5) over a range of CECs concentration (0-4000 μg/L). 10 μg/L of eight surrogate compounds (Table A-4) were added in each sample as internal standards.

The calibration curves were prepared by plotting the ratio of analyte to surrogate compound peak area as a function of analyte concentration.

B-2. Batch sorption experiment

A batch adsorption experiment was performed to evaluate the partitioning of all selected CECs to the activated carbon media. Pre-combusted amber glass bottles (500 mL) were filled with water (to avoid water-air partitioning) containing all selected CECs. Carbonate-buffered (0.1 M) ultrapure water (pH 8.5) was prepared and spiked with CECs at a concentration of 2 mg/L each. Activated carbon (AC) media was added into the bottles at various dosages to initiate the adsorption experiment. The same type of activated carbon as in the pilot-scale columns was used in this batch experiment (Calgon F300) but was first ground and sieved (65 mesh). The activated carbon dosage range (0–312.5 mg AC/L) was targeted so as to sorb 10%-90% of all CECs in order to generate sorption isotherms. In additional experiments, the water was dosed with Mississippi River NOM (obtained from the International Humic Substances Society) to a concentration of 5 mg C/L to investigate the effect of natural organic matter (NOM) on sorption capacities. The bottles were agitated at 120 rpm (Orbit Shaker-3520, Lab-line Instruments, Inc. Melrose Park, IL) at room temperature (~20 ^{o}C) until the adsorption equilibrium was reached (72 hours). Preliminary experiments were used to establish the sorption kinetics to determine the appropriate incubation time. The amount of CECs adsorbed was determined by measuring the initial and final concentration in the aqueous phase and computing the difference.

For the batch experiment, the higher concentration and a less complicated matrix permitted most CECs (except TCEP) to be quantified using a relatively simple high-performance liquid chromatography (HPLC) with ultraviolet–visible spectroscopy (UV-Vis). TCEP was quantified using LC-MS/MS as described above.

For HPLC-UV-Vis, a 4.6mm × 150mm Ascentis RP Amide column with 5 μm particle size (Supelco) was used to separate most CECs (except TCEP). A binary gradient consisting of 10 mM phosphate buffer in water with 10% (v/v) acetonitrile (A) and 100% acetonitrile (B) at a flow rate of 1 mL/min was used. The gradient program was as follows: 10% A held for 5 min, then gradually increased to 90% in 10 min and held at 90% for 5 min. The detection wavelength of UV-Vis was set at 220 nm.

B-3. Explanation for the $^{BV}/_{C_{inf}}$ term in the multiple linear regression analysis

The value of $^{BV}/_{C_{inf}}$ was used to represent q_c/q_e because $^{BV}/_{C_{inf}}$ can be considered proportional to q_c/q_e under the assumption that the difference between C_{eff_GAC} and C_{eff_Anth} is relatively constant:

$$q_c = \frac{\sum Q\left[\left(C_{inf} - C_{eff\_GAC}\right) - \left(C_{inf} - C_{eff\_Anth}\right)\right]t}{M_{GAC}}$$

$$= \sum \frac{Qt}{M_{GAC}}\left(C_{eff\_Anth} - C_{eff\_GAC}\right) = \frac{BV}{\rho_{GAC}}\left(C_{eff\_Anth} - C_{eff\_GAC}\right)_{average}$$

So: $\frac{q_c}{q_e} = \frac{\frac{BV}{\rho_{GAC}}(C_{eff\_Anth} - C_{eff\_GAC})_{average}}{K_d C_{inf}} = \frac{BV}{C_{inf}} \times \frac{(C_{eff\_Anth} - C_{eff\_GAC})_{average}}{\rho_{GAC} \times K_d}$, where ρ_{GAC} is the GAC media apparent density (g/L). If assume $(C_{eff\_Anth} - C_{eff\_GAC})_{average}$ is constant overtime, q_c/q_e is proportional to BV/C_{inf}.

The plots of q_c/q_e versus BV/C_{inf} of all selected CECs are in Figure B-4. The correlation coefficients from linear regression analysis (with intercepts fixed as zero) ranged from 0.985 to 0.997, suggesting that BV/C_{inf} was proportional to q_c/q_e.

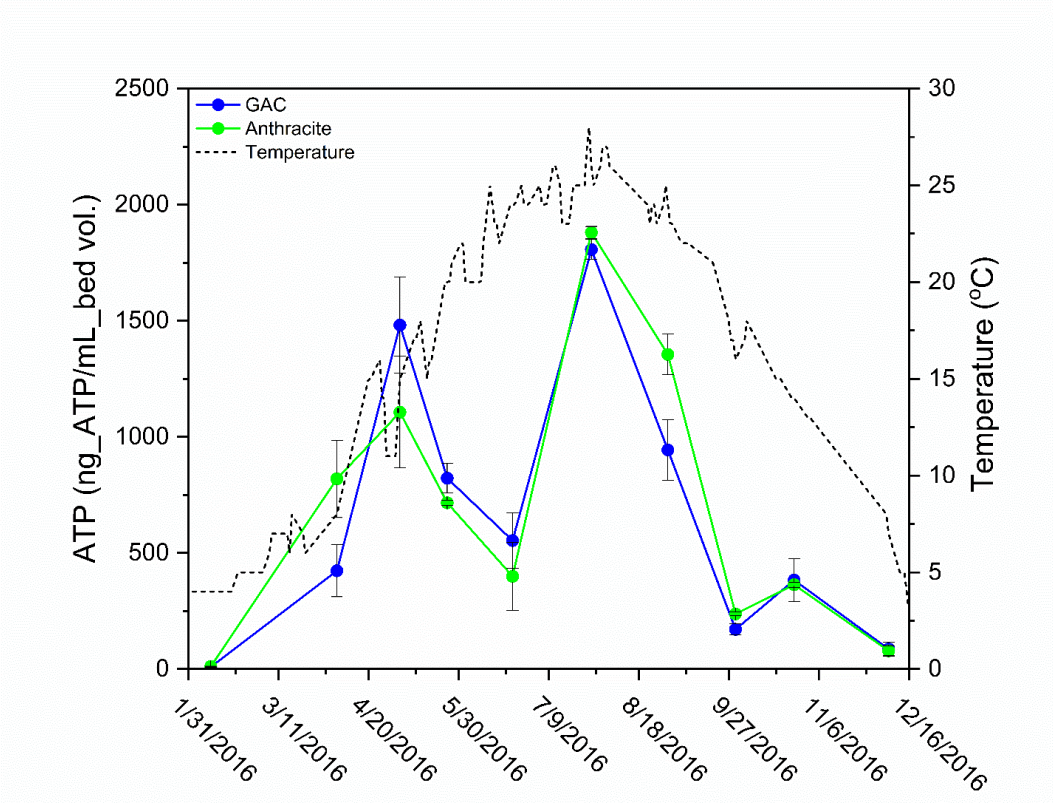

Figure B- 1. ATP concentrations at the top layer of the pilot-scale GAC-sand and anthracite-sand biofilter media as a function of time. The dashed line represents the influent water temperature.

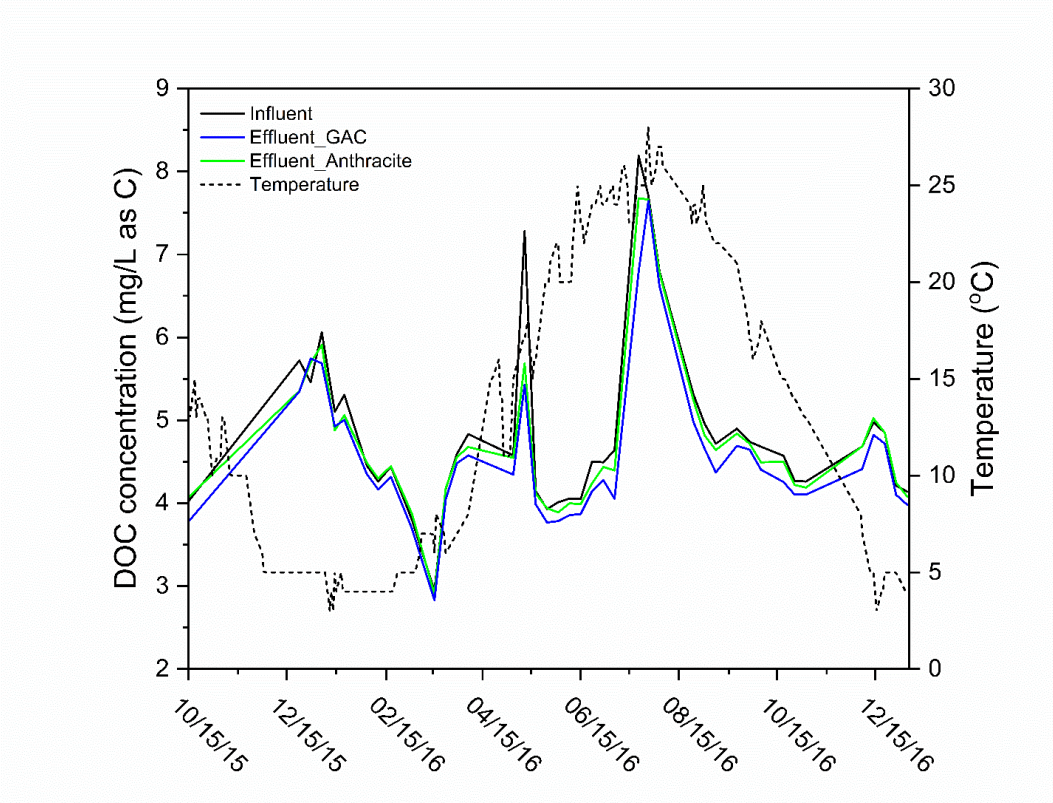

Figure B- 2. Mean DOC concentration in the filter influent (black), and filter effluent collected from the pilot-scale GAC-sand (blue) and anthracite-sand (green) biofilters. Dashed line represents the influent water temperature.

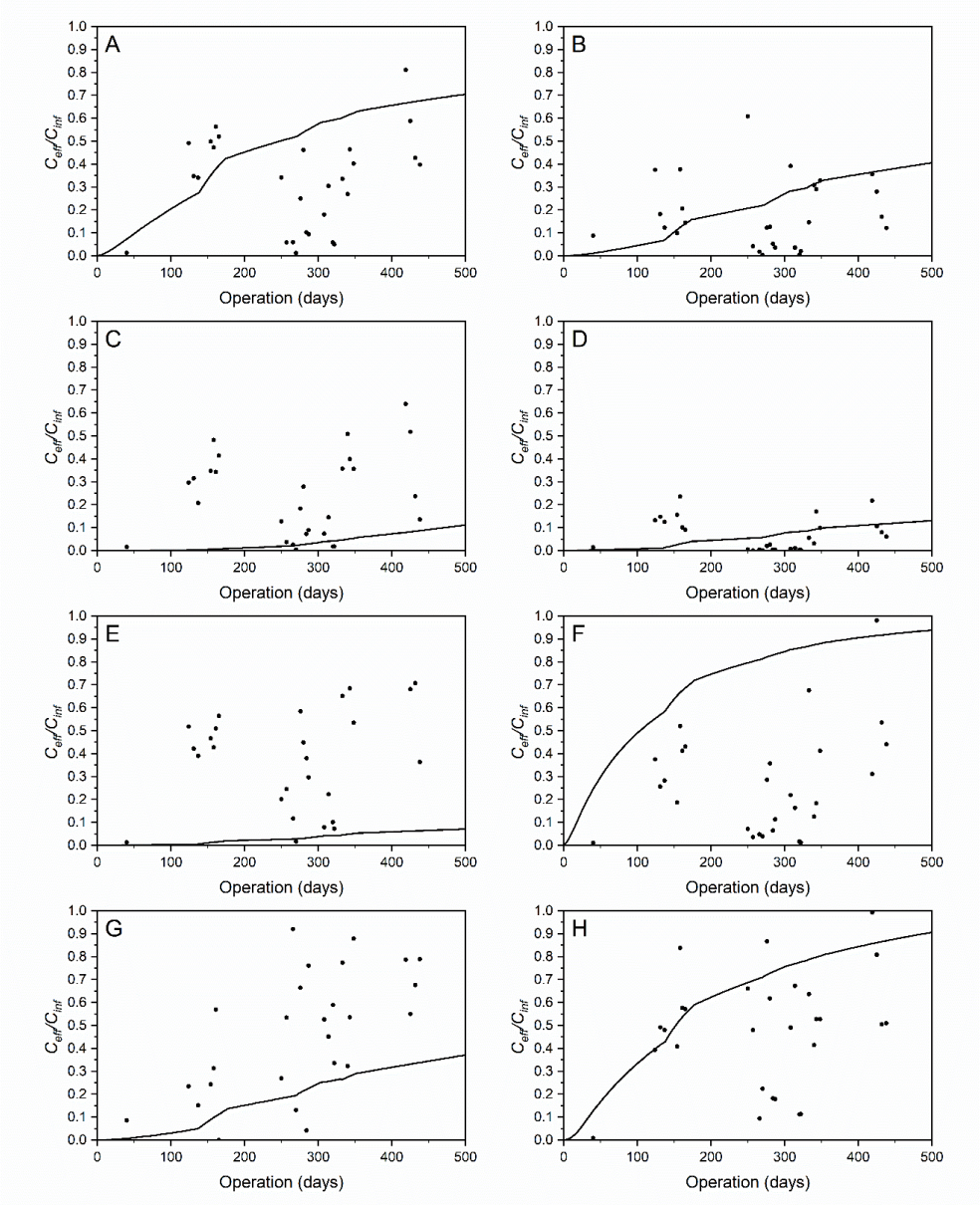

Figure B- 3. Comparison of predicted C_{eff}/C_{inf} values in GAC-sand biofilters using AdDesignS software versus the observed values for: (**A**) Atenolol, (**B**) Atrazine, (**C**) Carbamazepine, (**D**) Fluoxetine, (**E**) Gemfibrozil, (**F**) Metolachlor, (**G**) Sulfamethoxazole, (**H**) TCEP. Solid lines represent the model predictions and the points represent the observed values.

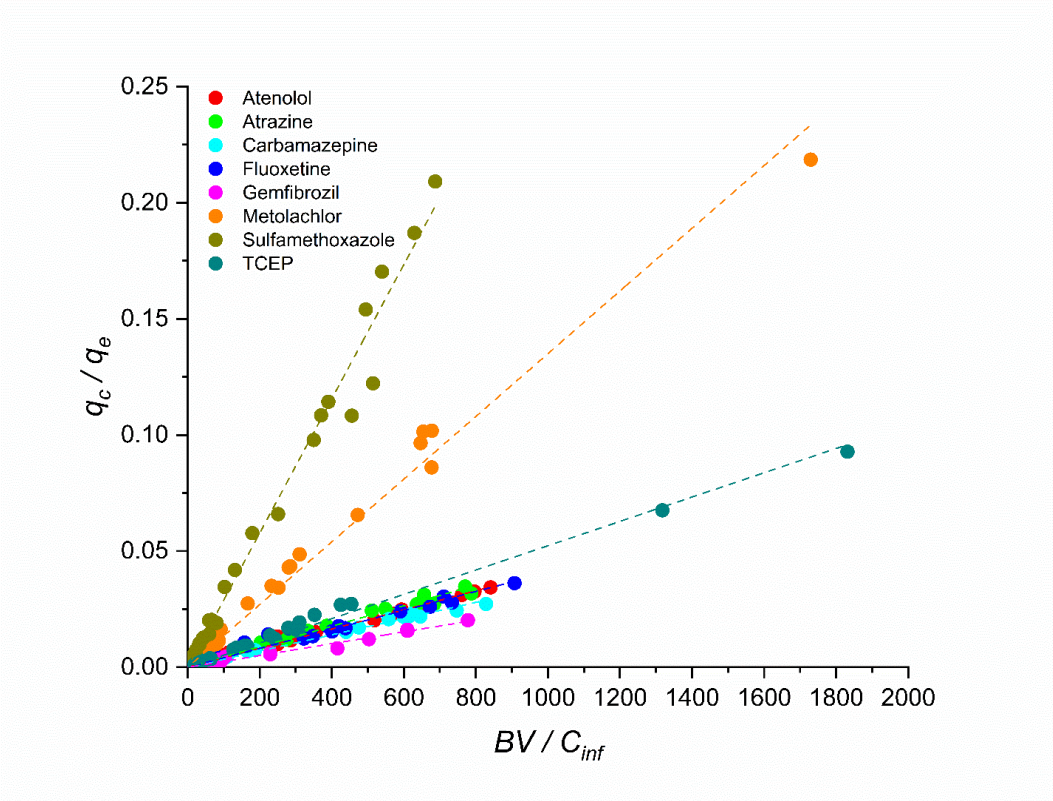

Figure B- 4. Plots of q_c/q_e versus $^{BV}/_{C_{inf}}$ for all selected CECs. The dashed lines represent the best fit linear regression lines (with intercept fixed at zero). The correlation coefficients (R^2) of linear regressions for atenolol, atrazine, carbamazepine, fluoxetine, gemfibrozil, metolachlor, sulfamethoxazole, and TCEP are 0.997, 0.994, 0.997, 0.985, 0.988, 0.992, 0.991, and 0.995, respectively.

Table B- 1. Influent concentrations (mean ± standard deviation) of all selected CECs. The schedule for different dosing is shown in Table 2-2.

	Low Influent Concentration (ng/L)	**High Influent Concentration (ng/L)**
Atenolol	167.4±77.2	1400.8±569.7
Atrazine	154.2±63.1	1380.5±468.8
Carbamazepine	144.1±50.8	1567.4±623.5
Fluoxetine	150.4±74.5	1296.7±468.8
Gemfibrozil	203.0±119.5	1836.4±816.8
Metolachlor	172.4±86.7	1633.1±837.2
Sulfamethoxazole	184.0±75.0	1773.3±683.0
TCEP	219.6±124.3	1264.3±617.3

Table B- 2. Chemical supplier and purity information of CECs and surrogate compounds.

	Supplier	**Purity/Chemical Grade**
Atenolol	Sigma	≥98% (TLC)
Atrazine	Sigma	analytical standard
Carbamazepine	Sigma	≥98% (HPLC)
Fluoxetine	United States Pharmacopeia	pharmaceutical primary standard
Gemfibrozil	Sigma-Aldrich	analytical standard
Metolachlor	Sigma-Aldrich	analytical standard
Sulfamethoxazole	Sigma-Aldrich	analytical standard
TCEP	Sigma-Aldrich	analytical standard
Atenolol-d_7	C/D/N Isotope, Inc	≥98%
Atrazine-d_5	C/D/N Isotope, Inc	≥98%
Carbamazepine-d_{10}	Cambridge Isotope Laboratories, Inc	≥98%
Fluoxetine-d_5	C/D/N Isotope, Inc	≥98%
Gemfibrozil-d_6	Toronto Research Chemicals	≥98%
Metolachlor-$^{13}C_6$	Cambridge Isotope Laboratories, Inc	≥98%
Sulfamethoxazole-d_4	Toronto Research Chemicals	≥98%
TCEP-d_{12}	Cambridge Isotope Laboratories, Inc	≥98%

Table B- 3. Isotherm (Langmuir) parameters (mean ± standard error) for all selected CECs obtained from fitting the batch test results for the following conditions: room temperature, 0.1 M carbonate buffered ultrapure water (pH 8.5) with 5 mg C/L of NOM.

	K_l (L/mg_CECs)	q_m (mg_CECs/mg_AC)	K_d (L/mg_AC)	R^2
Atenolol	94.5±192.7	0.028±0.002	2.65	0.55
Atrazine	84.7±29.6	0.022±0.001	1.86	0.97
Carbamazepine	79.1±147.9	0.035±0.002	2.77	0.69
Fluoxetine	45.3±36.4	0.075±0.005	3.40	0.88
Gemfibrozil	431.0±113.4	0.043±0.002	18.53	0.88
Metolachlor	32.8±17.1	0.062±0.006	2.04	0.78
Sulfamethoxazole	141.9±64.0	0.018±0.001	2.56	0.55
TCEP	16.9±5.9	0.028±0.003	0.47	0.70

Table B- 4. Compound-dependent parameters for LC-MS/MS analysis.

Selected CECs or surrogate compounds	Retention Time (min)	Precursor Ion m/z	Product Ion m/z	Collision Energy (eV)	LOD (ng/L)	LOQ (ng/L)
Atenolol	12.13	267.1	145.1	27	0.10	0.30
Atrazine	16.98	216.2	174.1	16	0.50	1.80
Carbamazepine	16.52	237.4	194	20	0.03	0.12
Fluoxetine	21.71	310.1	148.2	9	0.15	0.50
Gemfibrozil	4.98	249.1	121.1	-13	0.86	2.86
Metolachlor	18.33	284.0	175.5	25	0.26	0.86
Sulfamethoxazole	12.35	254.3	156	25	0.10	0.34
TCEP	16.05	284.9	125.1	20	1.46	4.86
Atenolol-d_7	12.04	274.2	145	26		
Atrazine-d_5	16.92	221.2	179	16		
Carbamazepine-d_{10}	16.47	247.4	204	20		
Fluoxetine-d_5	21.66	315.1	153.2	9		
Gemfibrozil-d_6	4.96	255.1	121.1	-13		
Metolachlor-$^{13}C_6$	18.39	290.0	181.5	25		
Sulfamethoxazole-d_4	12.30	258.3	160	25		
TCEP-d_{12}	16.06	297.0	130.1	20		

Appendix C

Supporting Information for Chapter 3

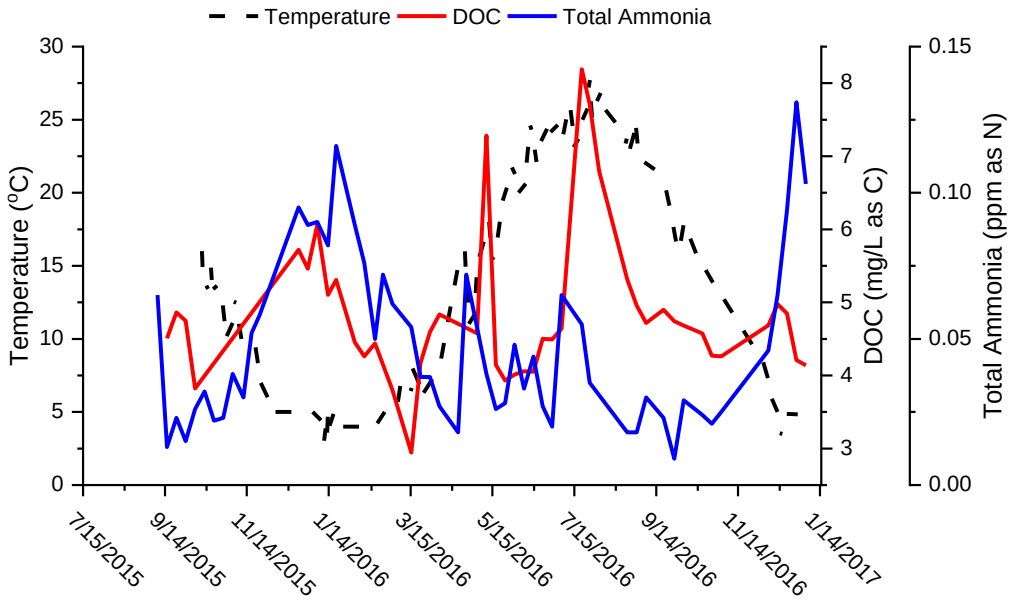

Figure C- 1. Temperature (black dashed line), DOC concentration (red line), and total ammonia concentration (blue line) in filter influent. Method detection limits for DOC and total ammonia are 0.004 mg C/L and 0.015 mg-N/L, respectively.

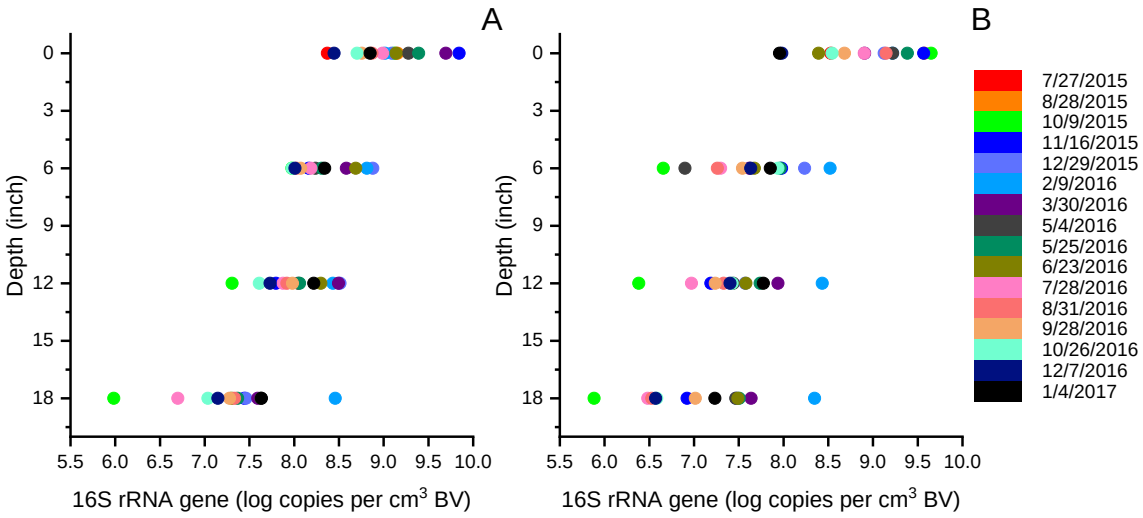

Figure C- 2. Mean 16S rRNA gene concentration at four difference depth of the pilot-scale GAC-sand (**A**, column #1-3) and anthracite-sand biofilters (**B**, column #4-6).

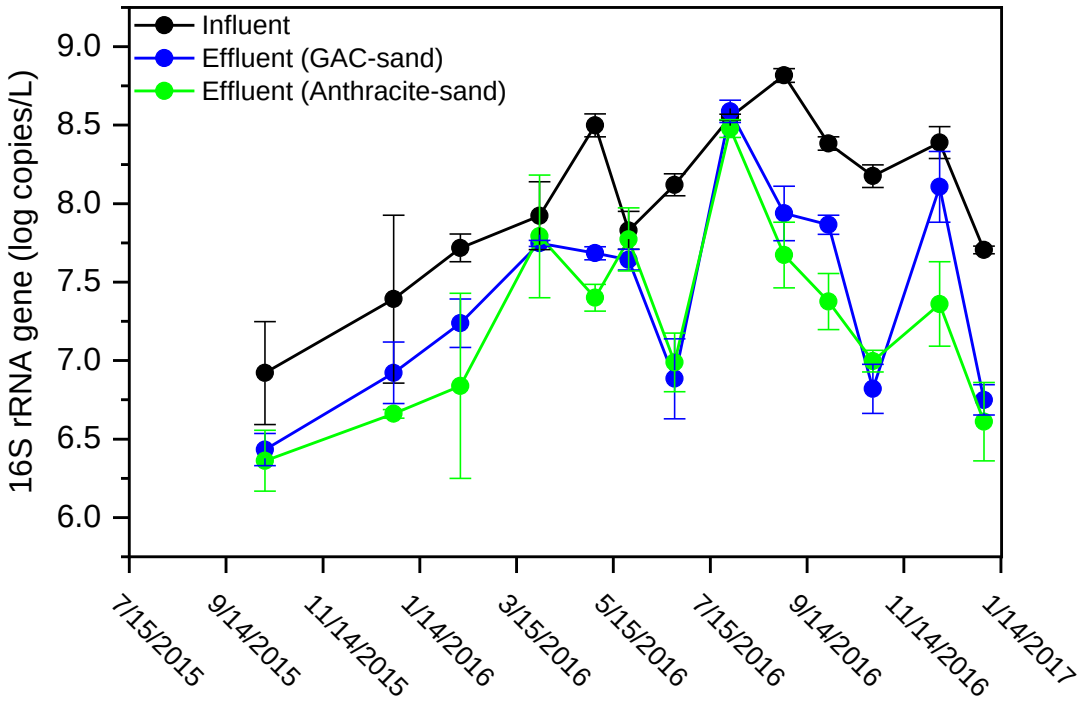

Figure C- 3. Mean 16S rRNA gene concentration in the filter influent (black), and filter effluent collected from the pilot-scale GAC-sand biofilters (blue; column #1-3) and anthracite-sand biofilters (green; column #4-6). Error bars represent one standard deviation from the mean of triplicate samples.

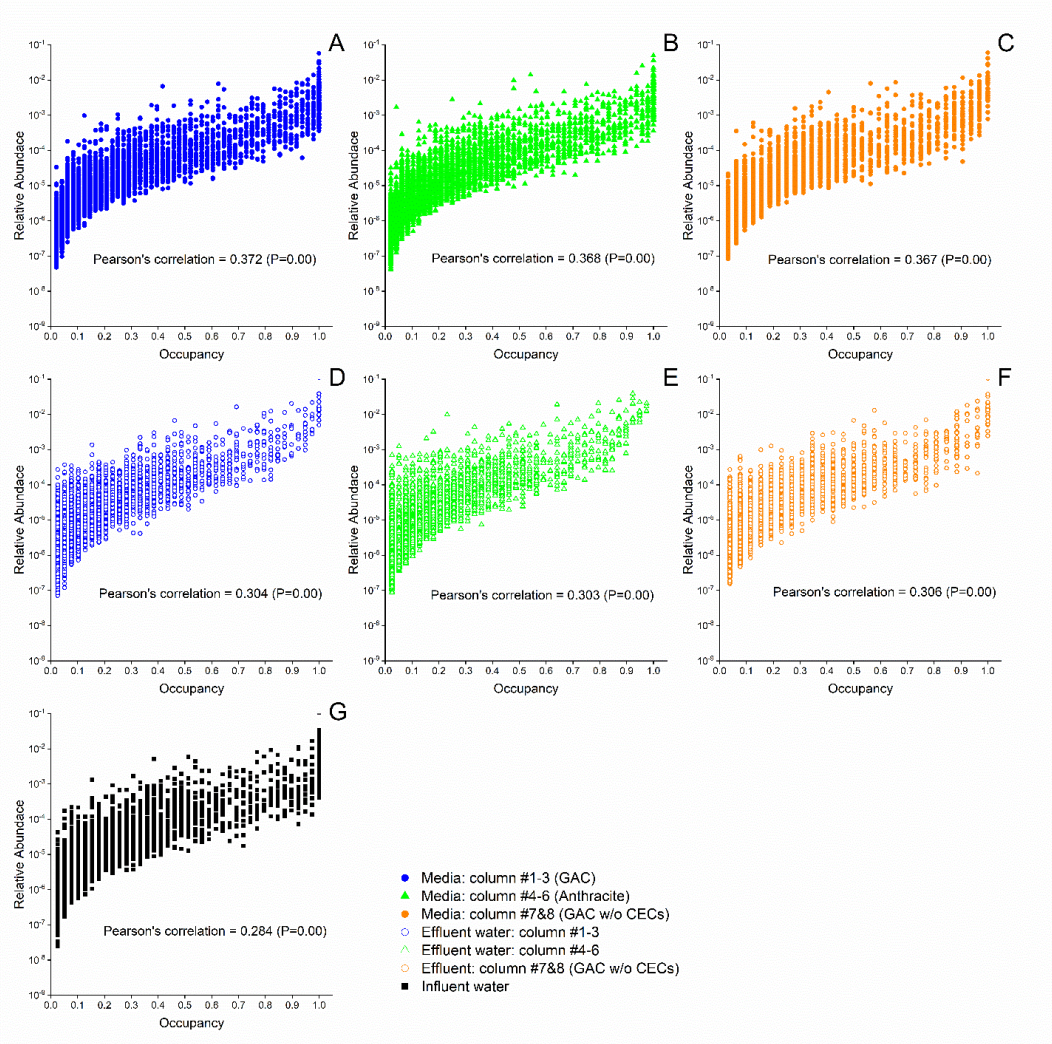

Figure C- 4. The relationships between ASV mean relative abundance and occupancy in media samples (**A-C**), filter influent samples (**D-F**) and filter effluent samples (**G**) collected from the pilot-scale biofilters.

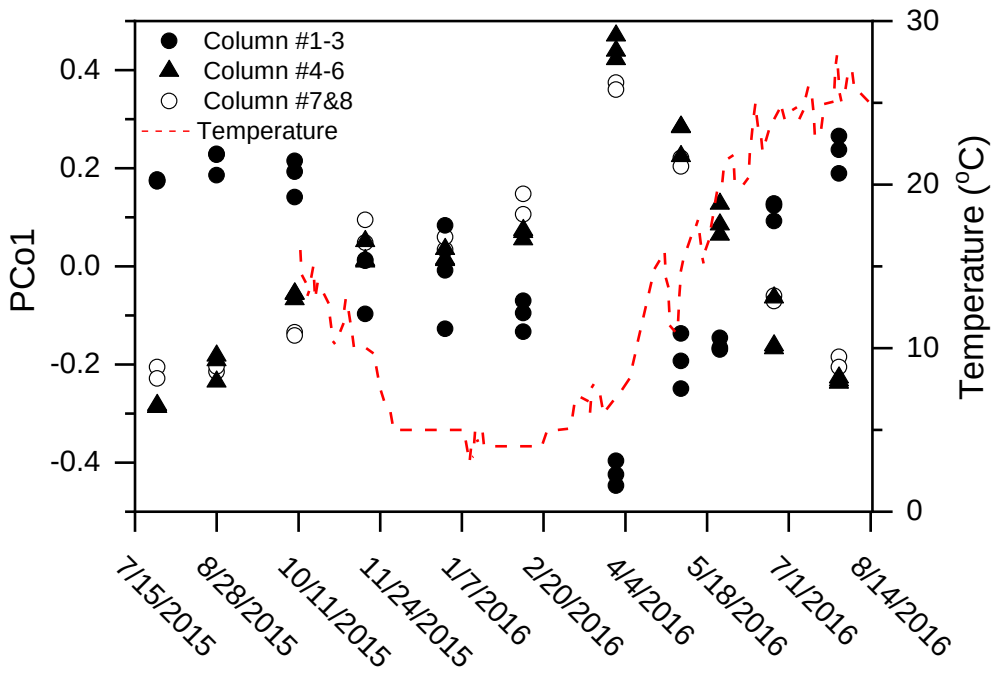

Figure C- 5. The first principal coordinates in the PCoA plots (Fig. 3-4) for the media samples collected the pilot-scale biofilters from 07/27/2015 to 07/28/2016. Red dashed line represents the influent water temperature.

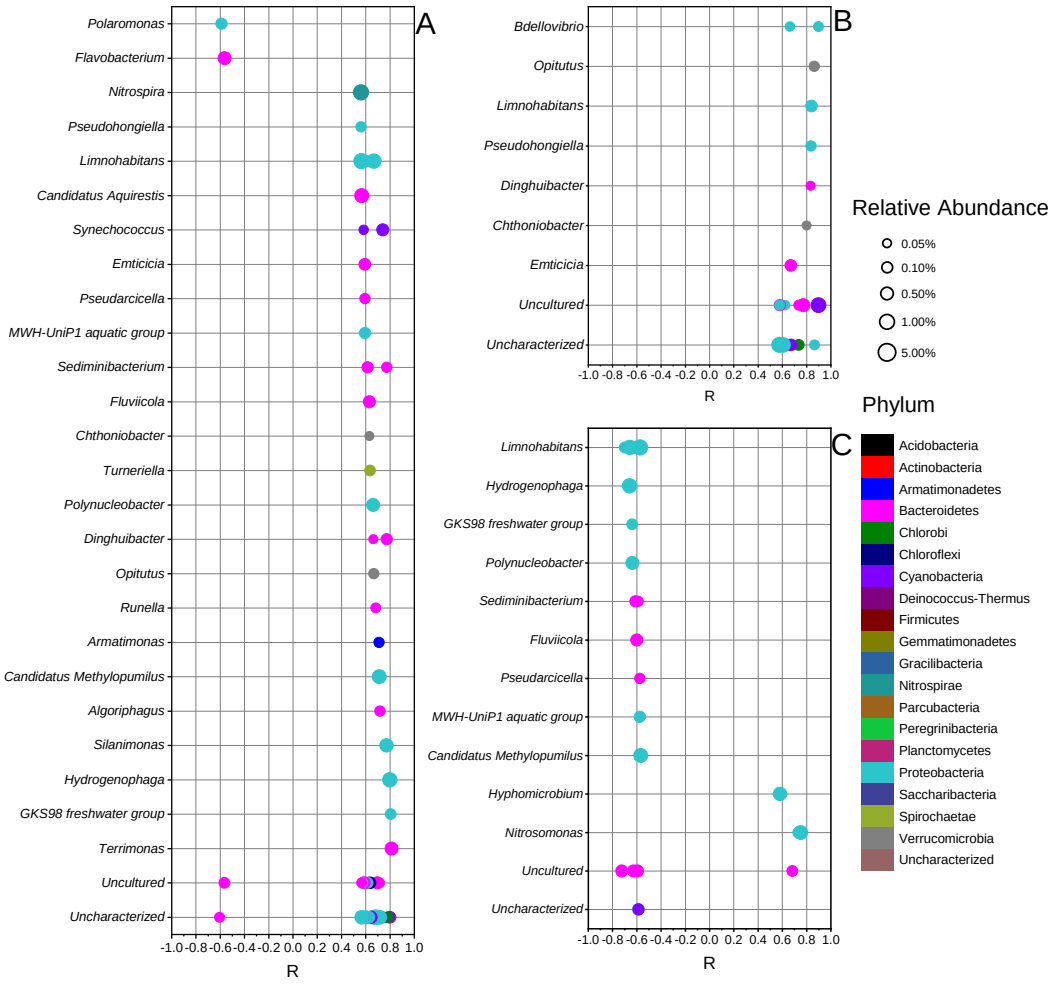

Figure C- 6. Pearson's correlations between ASV relative abundance in GAC media samples collected from the pilot-scale GAC-sand biofilters (column #1-3) and temperature (**A**), influent DOC concentration (**B**) and influent ammonia concentration (**C**). Only core ASVs with statistically significant correlations ($P_{adjusted}$ < 0.05, Pearson's correlation, Bonferroni correction) are listed here. X-axis is the Pearson's correlation coefficient. Y-axis is the taxonomic assignment of the ASVs. Color represents the Phylum of the ASV. The size of each dot represents the relative abundance.

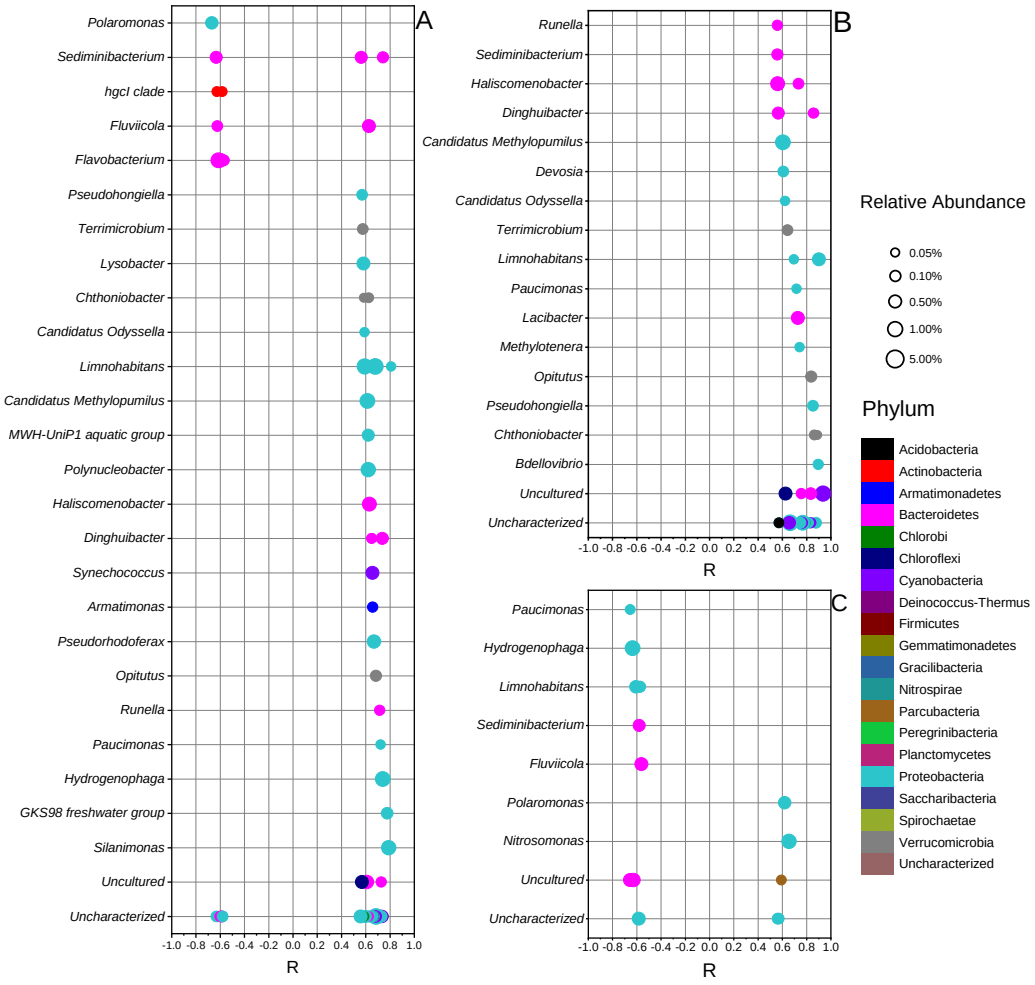

Figure C- 7. Pearson's correlation between ASV relative abundance in Anthracite media samples collected from column #4-6 and temperature (**A**), influent DOC concentration (**B**) and influent ammonia concentration (**C**). Only Core ASVs with statistically significant correlations ($P_{adjusted} < 0.05$, Pearson's correlation, Bonferroni correction) are listed here. X-axis is the Pearson's correlation coefficient. Y-axis is the taxonomic assignment of the ASVs. Color represents the Phylum of the ASV. The size of each dot represents the relative abundance.

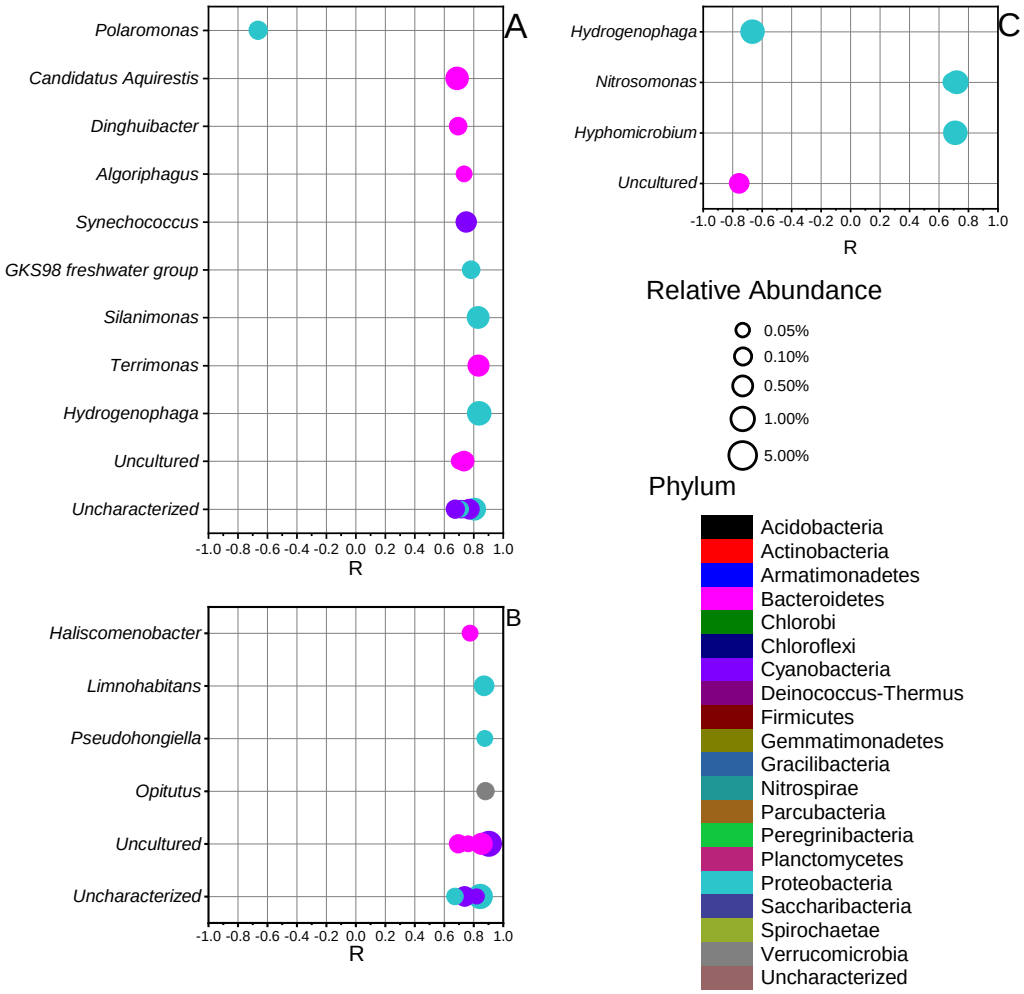

Figure C- 8. Pearson's correlation between ASV relative abundance in GAC media samples collected from the pilot-scale GAC-sand biofilters (column #7 and 8) and temperature (**A**), influent DOC concentration (**B**) and influent ammonia concentration (**C**). Only Core ASVs with statistically significant correlations ($P_{adjusted}$ < 0.05, Pearson's correlation, Bonferroni correction) are listed here. X-axis is the Pearson's correlation coefficient. Y-axis is the taxonomic assignment of the ASVs. Color represents the Phylum of the ASV. The size of each dot represents the relative abundance.

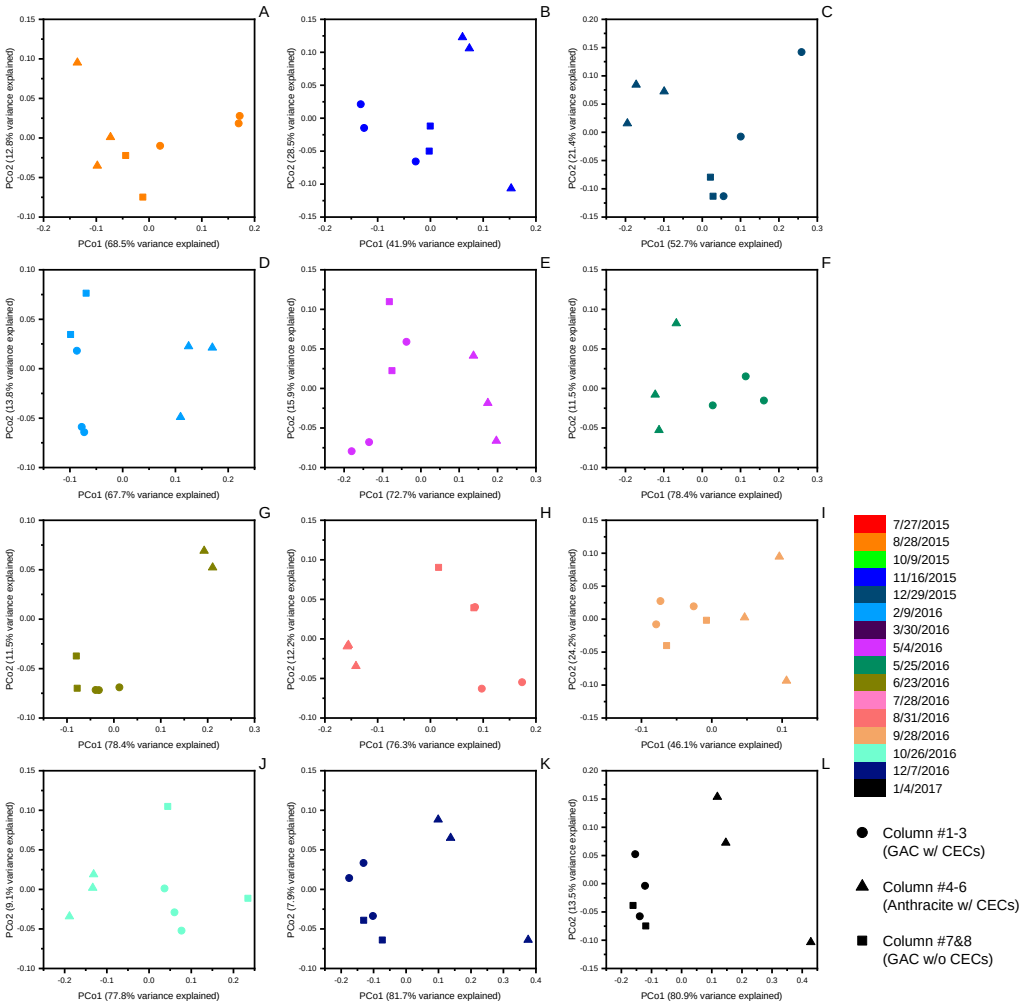

Figure C- 9. PCoA plots based on weighted UniFrac distance showing the effect of media type and CECs addition on the bacterial community in the pilot-scale biofilters. Different color represents the sampling date.

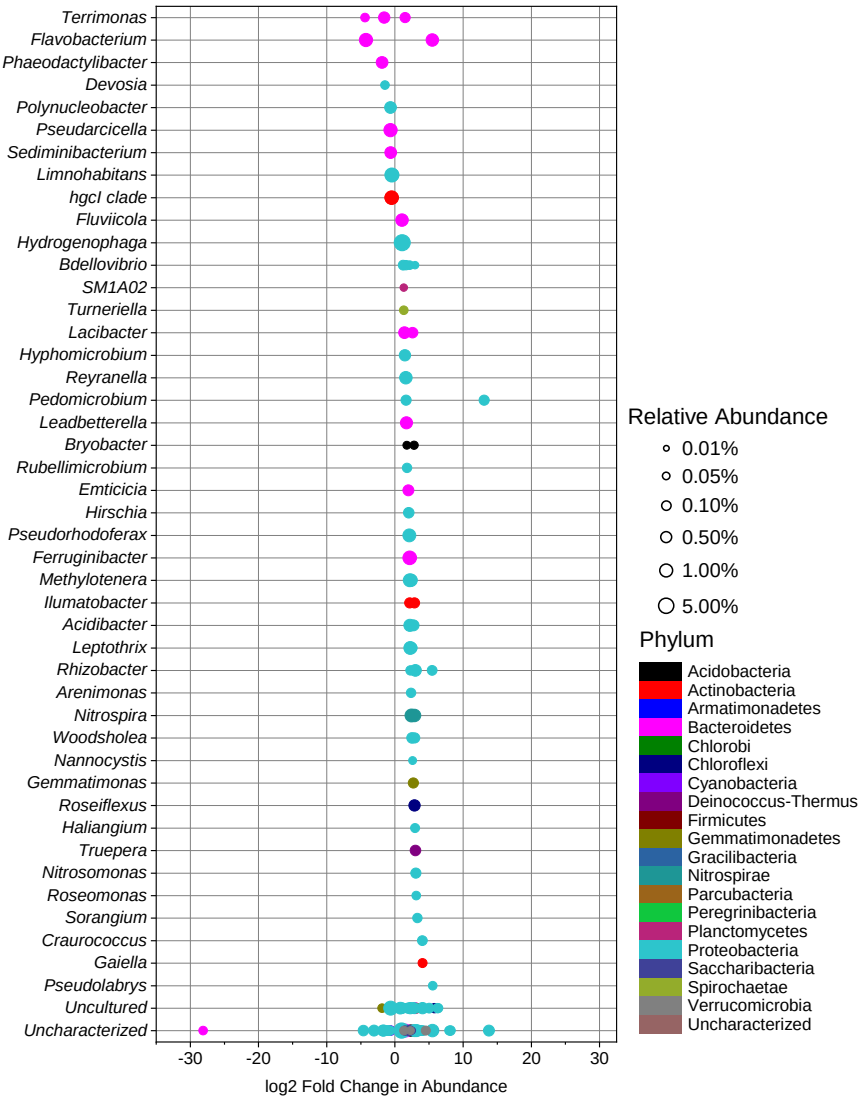

Figure C- 10. Differential abundance analysis of the ASV relative abundance on GAC media samples compared to anthracite media samples collected from the pilot-scale biofilters. Core ASVs with significant difference ($P_{walt\_adjusted}$<0.05) are grouped by genera in Y-axis. Color represents the Phylum of the ASV. The size of each dot represents the relative abundance.

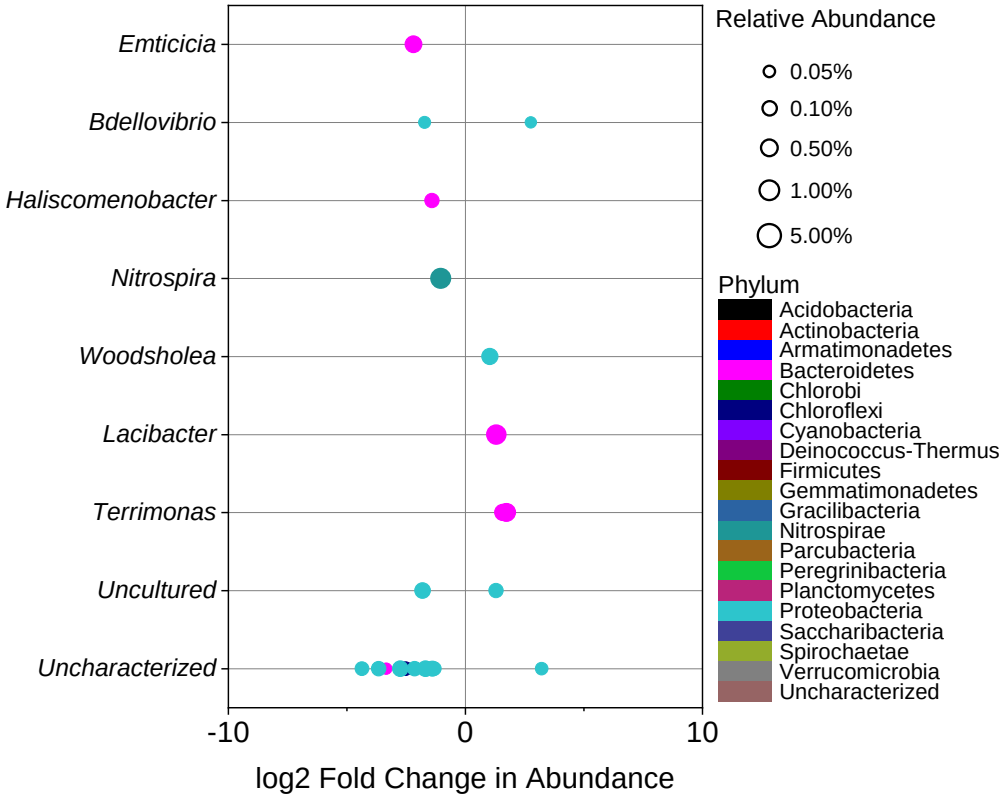

Figure C- 11. Differential abundance analysis of the ASV relative abundance in GAC media collected from the pilot-scale GAC-sand biofilters with continuous CECs addition (column #1-3) compared to those without continuous CECs addition (column #7&8). Core ASVs with significant difference ($P_{walt\_adjusted}$<0.05) are grouped by genera in Y-axis. Color represents the Phylum of the ASV. The size of each dot represents the relative abundance.

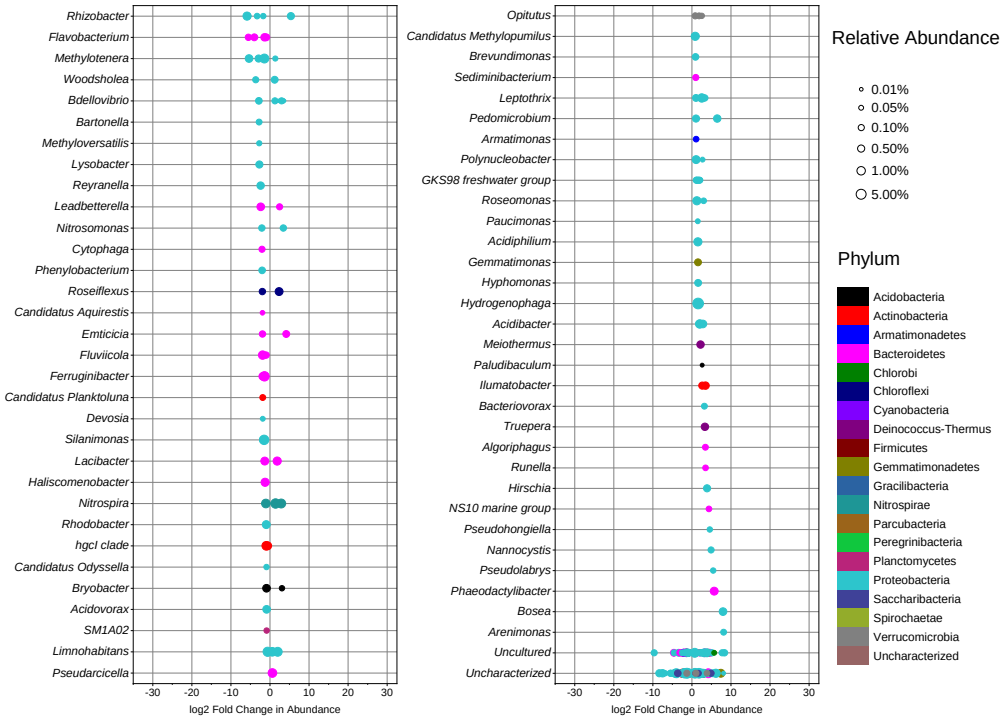

Figure C- 12. Differential abundance analysis of the ASVs relative abundance in the media samples collected from 08/15 to 01/16 (chloraminated backwash) compared to those collected from 08/16 to 01/17 (non-chloraminated backwash) in the pilot-scale biofilters. Core ASVs with significant difference ($P_{walt\_adjusted}$<0.05) are grouped by genera in Y-axis. Color represents the Phylum of the ASV. The size of each dot represents the relative abundance.

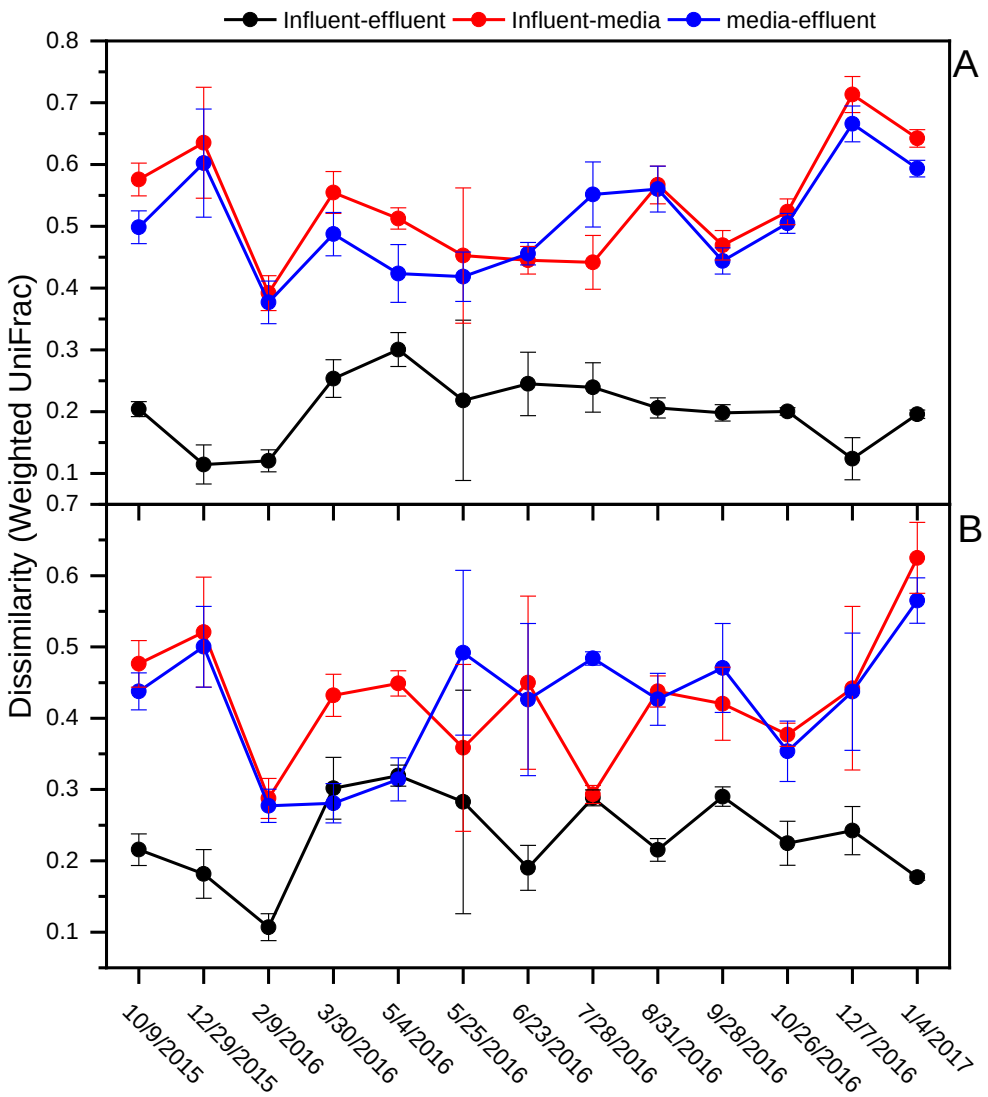

Figure C- 13. Weighted UniFrac distance of the bacterial community in filter influent, filter media, and filter effluent samples collected from the pilot-scale GAC-sand (**A**, column #1-3) and anthracite-sand biofilters (**B**, column #4-6). The error bars represent the standard deviation computed from 9 different UniFrac distance values (triplicate samples vs. triplicate samples).

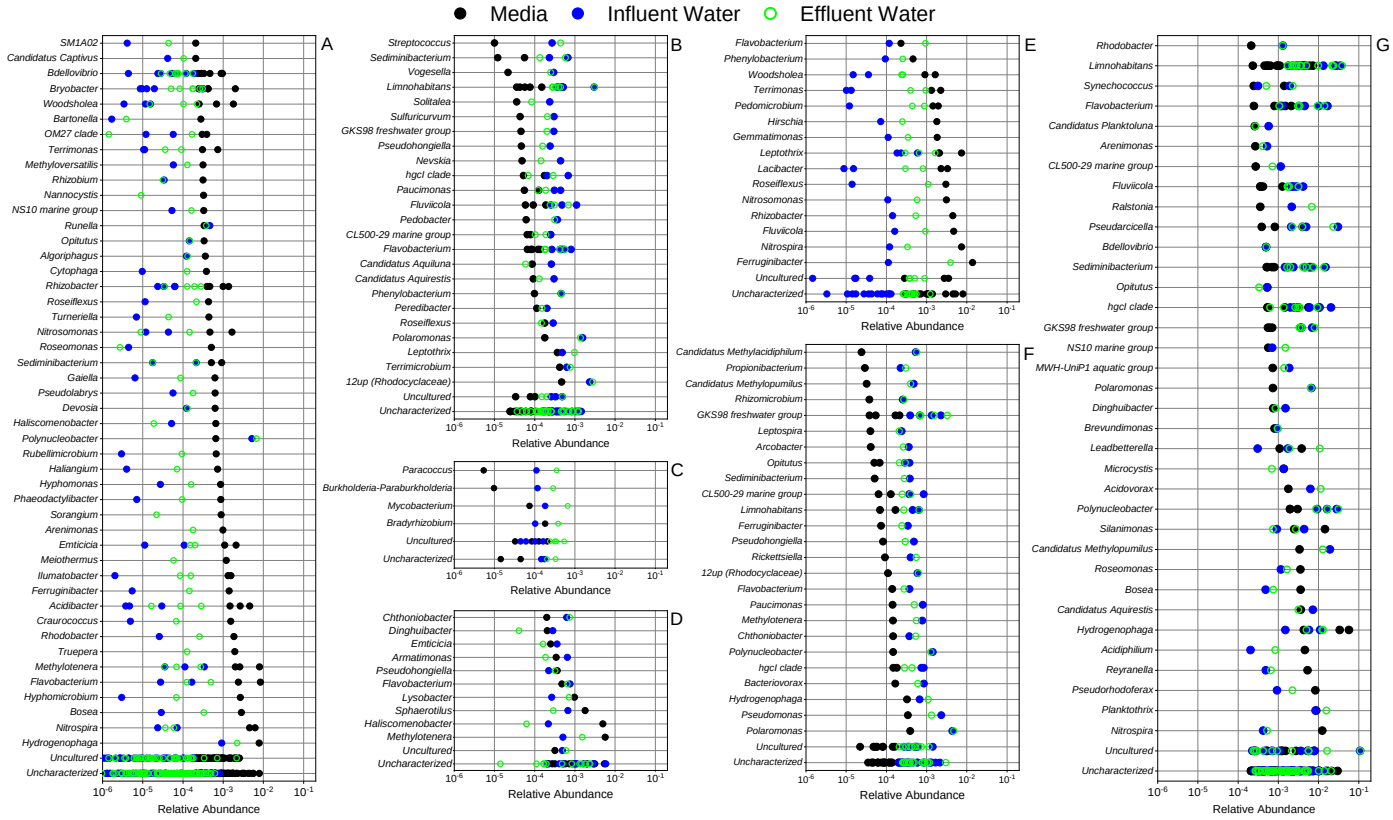

Figure C- 14. The mean relative abundance of all core ASVs in filter media (black), filter influent (blue) and filter effluent (green) samples that collected from the pilot-scale anthracite-sand biofilters (column #4-6). The ASVs are grouped as media specific (**A**), influent specific (**B**), effluent specific (**C**), influent-media shared (**D**), effluent-media shared (**E**), influent-effluent shared (**F**), and media-influent-effluent shared (**G**). The ASVs are grouped by genera in Y-axis.

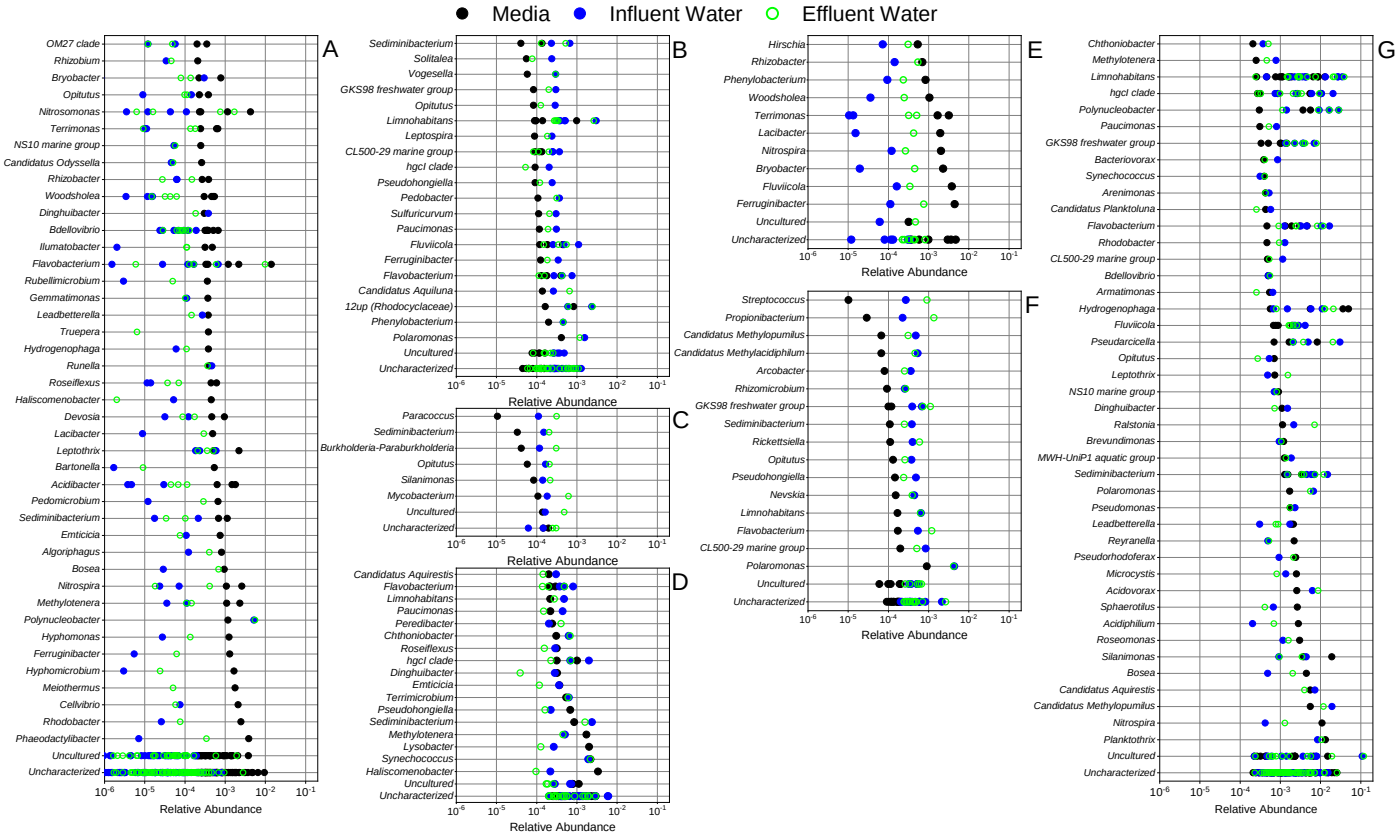

Figure C- 15. The mean relative abundance of all core 'media-influent-effluent shared' ASVs in filter media (black), filter influent (blue) and filter effluent (green) samples that collected from the pilot-scale GAC-sand (column #1-3) biofilters. The ASVs are grouped as media specific (**A**), influent specific (**B**), effluent specific (**C**), influent-media shared (**D**), effluent-media shared (**E**), influent-effluent shared (**F**), and media-influent-effluent shared (**G**). The ASVs are grouped by genera in Y-axis.

Table C- 1. Information of primer sequences and gBlock standards for 16S rRNA gene.

Gene	Primer sequence (5' → 3')
Primer	**F:** GAGAGGAAGGTCCCCCAC **R:**CGCTACTTGGCTGGTTCA G
gBlock	CGGTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGC ACAATGGGCGCAAGCCTGATGCAGCCATGCCGCGTGTATGA AGAAGGCCTTCGGGTTGTAAAGTACTTTCAGCGGGGAGGAA GGGAGTAAAGTTAATACCTTTGCTCATTGACGTTACCCGCAG AAGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATA CGGAGGGTGCAAGCGTTA

Table C- 2. Shannon index (Mean ± S.D., triplicate samples) for bacterial communities in media samples collected from top layer of the pilot-scale biofilters (GAC-sand biofilters: column #1-3; anthracite-sand biofilters: column #4-6) as measured by Illumina HiSeq analysis of 16S rRNA gene V3 region. (Sequence libraries were randomly trimmed down to 100,000 sequences per profile)

Date	GAC	Anthracite
7/27/2015	7.02±0.09	7.38±0.04
8/28/2015	7.89±0.07	7.40±0.14
10/9/2015	8.21±0.09	7.80±0.09
11/16/2015	7.20±0.08	6.99±0.14
12/29/2015	6.82±0.25	7.36±0.25
2/9/2016	7.49±0.13	7.59±0.06
3/30/2016	5.57±0.22	5.84±0.18
5/4/2016	6.56±0.19	5.63±0.26
5/25/2016	6.61±0.10	6.59±0.20
6/23/2016	7.70±0.02	7.20±0.36
7/28/2016	8.24±0.07	8.07±0.05
8/31/2016	8.34±0.03	8.09±0.05
9/28/2016	8.26±0.03	8.16±0.05
10/26/2016	7.63±0.03	7.19±0.12
12/7/2016	7.91±0.04	8.13±0.04
1/4/2017	6.61±0.06	5.69±0.30

Table C- 3. Shannon index (Mean ± S.D., triplicate samples) for bacterial communities in media samples collected from four different depth in the biofilters (GAC-sand biofilters: column #1-3; anthracite-sand biofilters: column #4-6) on 07/28/2016 and 09/28/2016, as measured by Illumina HiSeq analysis of 16S rRNA gene V3 region. (Sequence libraries were randomly trimmed down to 100,000 sequences per profile)

	GAC	Anthracite
7/28/2016		
0 in.	8.24±0.07	8.07±0.05
6 in.	7.92±0.13	7.92±0.13
12 in.	7.54±0.27	7.54±0.27
18 in.	6.91±0.67	6.44±0.46
9/28/2016		
0 in.	8.26±0.03	8.16±0.05
6 in.	7.83±0.10	8.61±0.09
12 in.	8.00±0.39	7.87±0.56
18 in.	7.71±0.56	7.67±0.82

Table C- 4. Shannon index (Mean ± S.D. (sample number)) for bacterial communities in the water samples collected from the pilot-scale biofilters as measured by Illumina HiSeq analysis of 16S rRNA gene V3 region. (Shannon index values for the samples with less than 100,000 high quality sequencing reads were not calculated).

Date	Influent	Effluent: GAC-sand	Effluent: Anthracite-sand
10/9/2015	6.25±0.04(3)	6.94±0.00(1)	6.21±0.00(1)
12/29/2015	6.97±0.22(2)	6.73±0.06(2)	6.74±0.05(3)
2/9/2016	6.85±0.06(3)	6.79±0.05(3)	6.81±0.09(2)
3/30/2016	7.46±0.08(3)	7.22±0.05(3)	6.44±0.18(3)
5/4/2016	7.07±0.06(3)	6.82±0.14(3)	6.62±0.12(3)
5/25/2016	6.94±0.14(3)	6.83±0.18(3)	7.06±0.05(3)
6/23/2016	6.33±0.32(3)	6.17±0.33(3)	6.04±0.13(2)
7/28/2016	7.39±0.15(3)	6.46±0.20(3)	6.21±0.02(3)
8/31/2016	6.97±0.27(3)	7.06±0.15(3)	7.11±0.14(3)
9/28/2016	7.11±0.08(3)	7.07±0.02(3)	6.73±0.11(3)
10/26/2016	6.53±0.12(3)	6.27±0.00(1)	5.99±0.25(2)
12/7/2016	7.50±0.39(3)	7.50±0.05(3)	7.16±0.13(3)
1/4/2017	5.92±0.35(3)	5.91±0.00(1)	5.78±0.00(1)

Table C- 5. Results from PERMANOVA tests (adonis) employed to test effect of temperature, DOC and ammonia on the bacterial community on the GAC and anthracite media samples collected from the pilot-scale biofilters.

	R^2_{adonis}	P_{adonis}	$P_{betadisper}$
GAC			
Temperature	0.249	0.001	0.001
DOC	0.055	0.049	0.923
Ammonia	0.065	0.024	0.001
Residual	0.631		
Anthracite			
Temperature	0.271	0.001	0.001
DOC	0.079	0.003	0.855
Ammonia	0.113	0.004	0.255
Residual	0.538		

Table C- 6. Results from PERMANOVA tests (adonis) employed to test statistically significant difference of microbial communities between the pilot-scale GAC-sand and anthracite-sand biofilters.

	Pairwise adonis			Pairwise betadisper	
	R^2_{adonis}	P_{adonis}	$P_{adonis\_adjusted}$	$P_{wilcoxon}$	$P_{wilcoxon\_adjusted}$
GAC vs. anthracite (all samples)	0.092	0.001		0.247	
GAC vs. anthracite (each sampling date)					
7/27/2015	0.728	0.140	0.171	1.000	1.000
8/28/2015	0.528	0.080	0.171	0.800	1.000
10/9/2015	0.788	0.100	0.171	0.800	1.000
11/16/2015	0.504	0.110	0.171	1.000	1.000
12/29/2015	0.370	0.100	0.171	1.000	1.000
2/9/2016	0.836	0.060	0.171	1.000	1.000
3/30/2016	0.931	0.090	0.171	1.000	1.000
5/4/2016	0.755	0.090	0.171	1.000	1.000
5/25/2016	0.711	0.130	0.171	1.000	1.000
6/23/2016	0.478	0.150	0.171	1.000	1.000
7/28/2016	0.819	0.170	0.181	1.000	1.000
8/31/2016	0.902	0.120	0.171	0.667	1.000
9/28/2016	0.309	0.310	0.310	1.000	1.000
10/26/2016	0.893	0.100	0.171	1.000	1.000
12/7/2016	0.617	0.110	0.171	1.000	1.000
1/4/2017	0.631	0.110	0.171	1.000	1.000

Table C- 7. Results from PERMANOVA tests (adonis) employed to test statistically significant difference of microbial communities in the pilot-scale GAC-sand biofilters with and without continuous CECs addition.

	Pairwise adonis			Pairwise betadisper	
	R^2_{adonis}	p_{adonis}	$p_{adonis\_adjusted}$	$p_{wilcoxon}$	$p_{wilcoxon\_adjusted}$
GAC w/ CECs vs. GAC w/o CECs (all samples)	0.004	0.879		1.000	
GAC w/ CECs vs. GAC w/o CECs (each sampling date)					
7/27/2015	0.306	0.310	0.495	1.000	1.000
8/28/2015	0.522	0.190	0.471	0.800	1.000
10/9/2015	0.211	0.500	0.600	0.800	1.000
11/16/2015	0.492	0.200	0.471	1.000	1.000
12/29/2015	0.227	0.580	0.600	1.000	1.000
2/9/2016	0.312	0.320	0.495	1.000	1.000
3/30/2016	0.931	0.090	0.471	1.000	1.000
5/4/2016	0.185	0.520	0.600	1.000	1.000
5/25/2016					
6/23/2016	0.481	0.100	0.471	1.000	1.000
7/28/2016	0.444	0.220	0.471	1.000	1.000
8/31/2016	0.439	0.190	0.471	0.800	1.000
9/28/2016	0.164	0.600	0.600	1.000	1.000
10/26/2016	0.352	0.220	0.471	1.000	1.000
12/7/2016	0.321	0.410	0.559	1.000	1.000
1/4/2017	0.318	0.330	0.495	1.000	1.000

Table C- 8. Results from PERMANOVA tests (adonis) employed to test statistically significant difference of microbial communities in media samples collected from four different depth (0 in., 6 in., 12 in., and 18 in.) in the pilot-scale GAC-sand (column #1-3) and anthracite-sand (column #4-6) biofilters.

Pairwise Comparison	Pairwise adonis			Pairwise betadisper	
	R^2_{adonis}	P_{adonis}	$P_{adonis\_adjusted}$	$P_{wilcoxon}$	$P_{wilcoxon\_adjusted}$
GAC-sand biofilters					
0" vs. 6 "	0.899	0.002	**0.004**	0.729	0.787
0" vs 12"	0.821	0.002	**0.004**	0.170	0.509
0" vs. 18"	0.629	0.001	**0.004**	0.514	0.771
6" vs. 12"	0.369	0.008	**0.010**	0.170	0.509
6" vs. 18"	0.585	0.005	**0.008**	0.302	0.603
12" vs. 18"	0.334	0.042	**0.042**	0.787	0.787
Anthracite-sand biofilters					
0" vs. 6 "	0.819	0.006	**0.012**	0.964	0.989
0" vs 12"	0.837	0.001	**0.006**	0.514	0.989
0" vs. 18"	0.712	0.004	**0.012**	0.514	0.989
6" vs. 12"	0.092	0.357	0.357	0.471	0.989
6" vs. 18"	0.186	0.156	0.234	0.729	0.989
12" vs. 18"	0.128	0.300	0.357	0.989	0.989

Table C- 9. Results from PERMANOVA tests (adonis) employed to test statistically significant difference in the bacterial communities in filter influent, filter media, and filter effluent collected from the pilot-scale GAC-sand (column #1-3) and anthracite-sand (column #4-6) biofilters.

Pairwise Comparison	Pairwise adonis			Pairwise betadisper	
	R^2_{adonis}	p_{adonis}	$p_{adonis\_adjusted}$	$p_{wilcoxon}$	$p_{wilcoxon\_adjusted}$
GAC media vs. GAC Influent	0.600	0.001	**0.002**	0.038	0.153
GAC media vs. GAC effluent	0.510	0.001	**0.002**	1.000	1.000
GAC Influent vs. GAC effluent	0.022	0.178	0.203	0.216	0.395
Anthracite media vs. Anthracite influent	0.427	0.001	**0.002**	0.005	0.038
Anthracite media vs. Anthracite effluent	0.311	0.001	**0.002**	0.465	0.620
Anthracite influent vs. Anthracite effluent	0.041	0.049	0.065	0.216	0.395

Appendix D

Supporting Information for Chapter 4

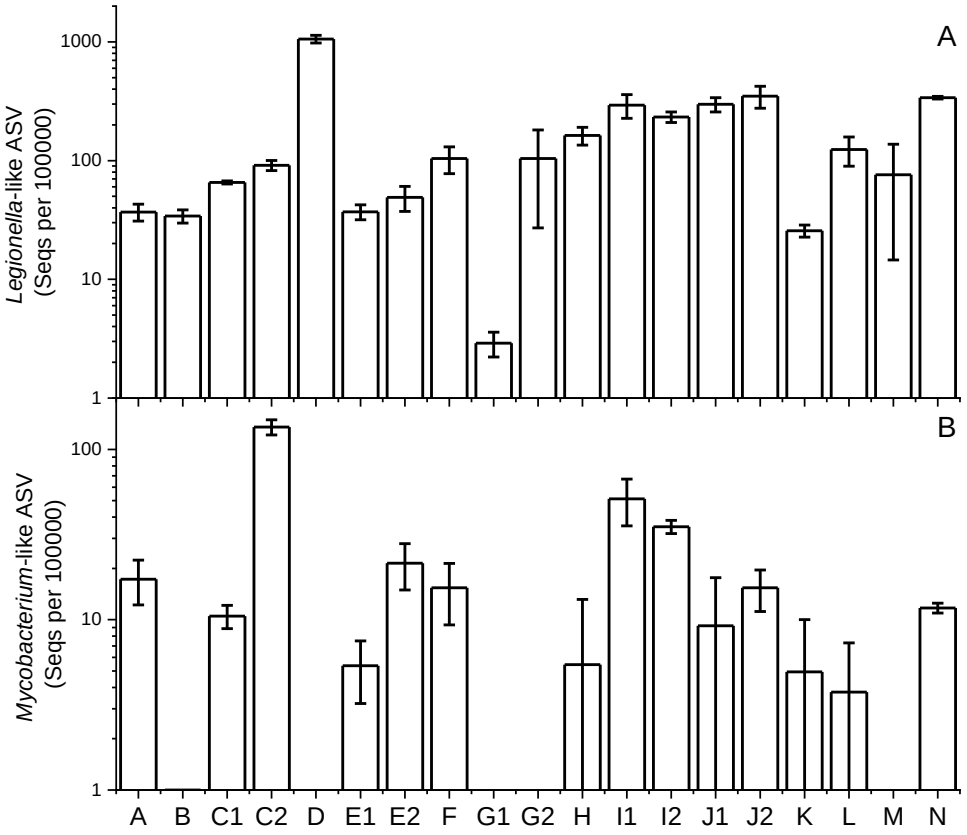

Figure D- 1. Relative abundance (in sequences per 100,000 reads) of *Legionella*-like ASV (**A**) and *Mycobacterium*-like ASV (**B**) on the full-scale biofilter media samples according to DNA sequencing results.

Table D- 1. Contribution of 26 most abundant genera to the microbial community composition in the full-scale biofilters. The abundant genera are those that comprised greater than 5% of total sequences in at least 3 of 57 samples in total that were collected in this investigation.

Genera	Mean relative abundance	# of samples with abundance > 5%	# of samples with detection
(Unknown) Comamonadaceae	6.60%	25	56
(Unknown) Sphingomonadales	5.43%	23	57
Bradyrhizobium	3.71%	16	56
(Unknown) Blastocatellaceae	3.56%	16	57
Nitrospira	2.81%	8	56
Hydrogenophaga	2.68%	6	40
(Uncultured) Deinococcaceae	2.48%	4	57
Leptothrix	2.26%	6	45
(Unknown) Hyphomicrobiaceae	2.12%	9	56
(Unknown) Rhodobacteraceae	1.96%	8	45
Acidovorax	1.87%	9	7
(Uncultured) Acetobacteraceae	1.82%	3	57
(Uncultured) Alcaligenaceae	1.51%	4	50
Blastocatella	1.39%	3	29
Hyphomicrobium	1.30%	3	28
Terrimonas	1.29%	5	42
(Unknown) Acetobacteraceae	1.29%	5	56
Bryobacter	1.25%	3	51
(Uncultured)Saprospiraceae	1.24%	3	56
(Unknown) Gallionellaceae	1.19%	6	26
(Uncultured) Methylobacteriaceae	1.13%	3	35
(Unknown) Betaproteobacteria	0.91%	4	9
Bdellovibrio	0.82%	4	11
(Unknown) Betaproteobacteria TRA3-20	0.59%	3	12
Silanimonas	0.44%	3	56
(Unknown) Gammaproteobacteria	0.35%	3	57

Table D- 2. Results from Mantel's tests employed to test correlations between geographic distance and water quality parameters (pH, temperature, and influent DOC concentration).

	R_{mantel}	P_{mantel}
pH	0.206	0.002
Temperature	0.107	0.001
Influent DOC concentration	0.190	0.003

Appendix E

Supporting Information for Chapter 5

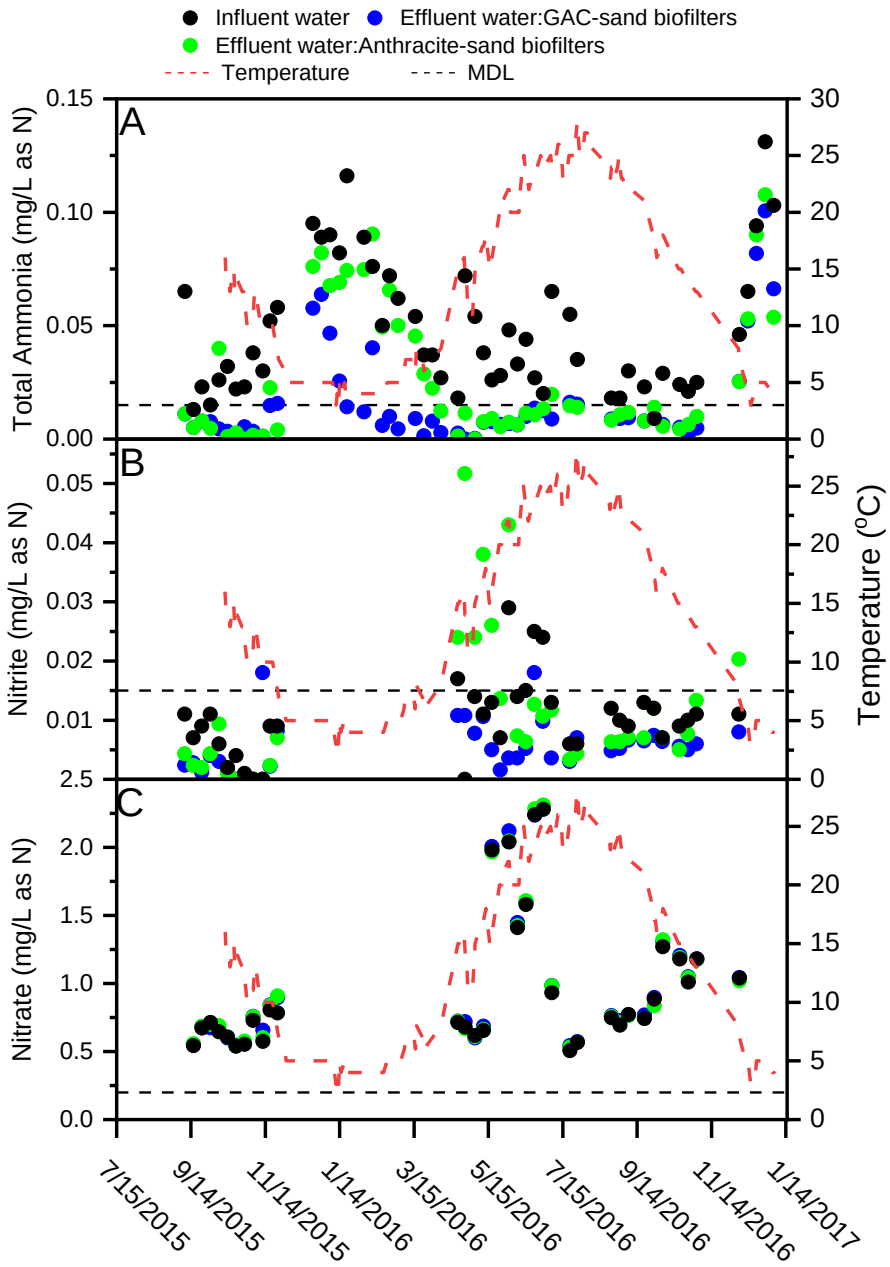

Figure E- 1. Concentrations of ammonia(**A**), nitrite (**B**) and nitrate (**C**) in the water samples collected from the pilot-scale biofilters.

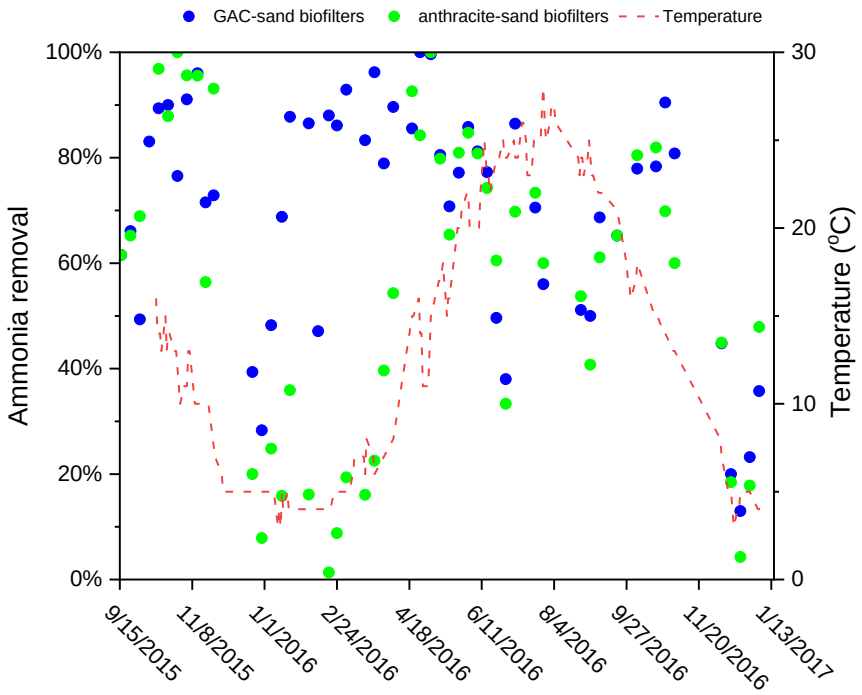

Figure E- 2. Ammonia removal in the pilot-scale GAC-sand and anthracite-sand biofilters. The red dashed line represents the water temperature.

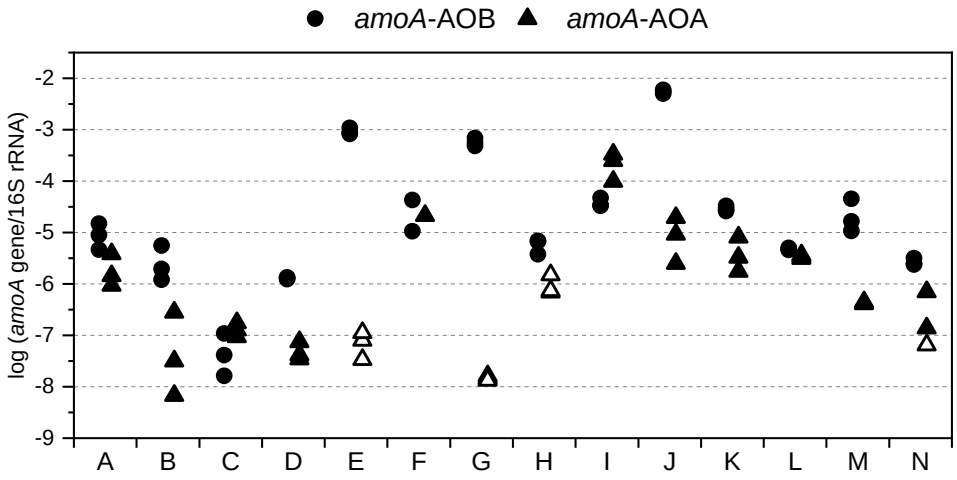

Figure E- 3. Ratio of *amoA* gene concentration to 16S rRNA gene concentration on the full-scale biofilter media samples. X-axis is the biofilter ID. Open symbol indicates that the *amoA* gene concentration is below the CL.

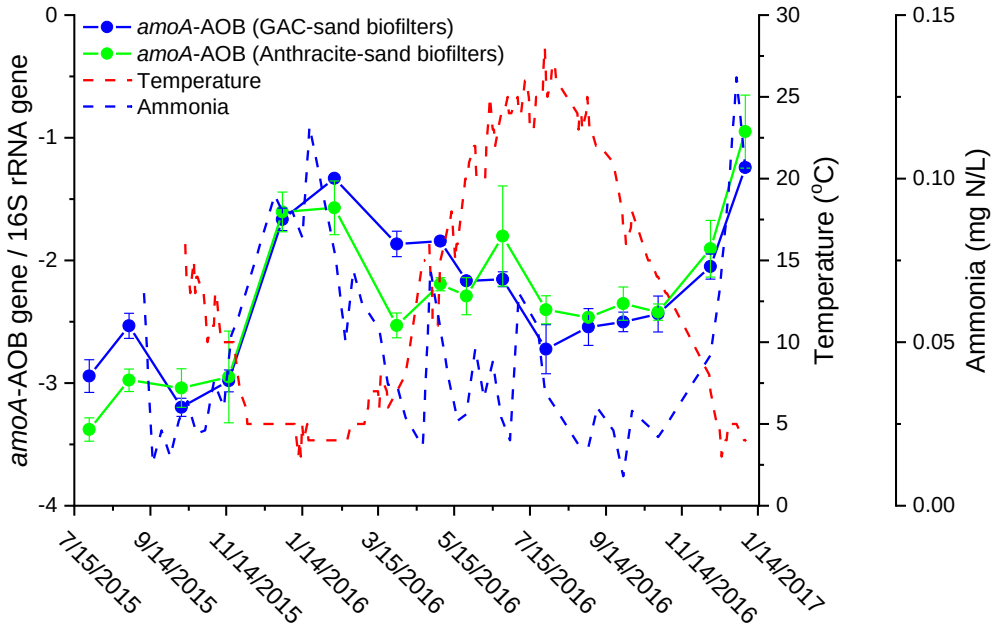

Figure E- 4. Ratio of *amoA* gene concentration to 16S rRNA gene concentration on media samples collected from pilot-scale GAC-sand (blue solid line) and anthracite-sand (green solid line) biofilters. Red dashed line represents the water temperature. Black dashed line represents the influent ammonia concentration. All gene concentrations are above censoring limits. Error bars represent the standard deviation computed from triplicate samples.

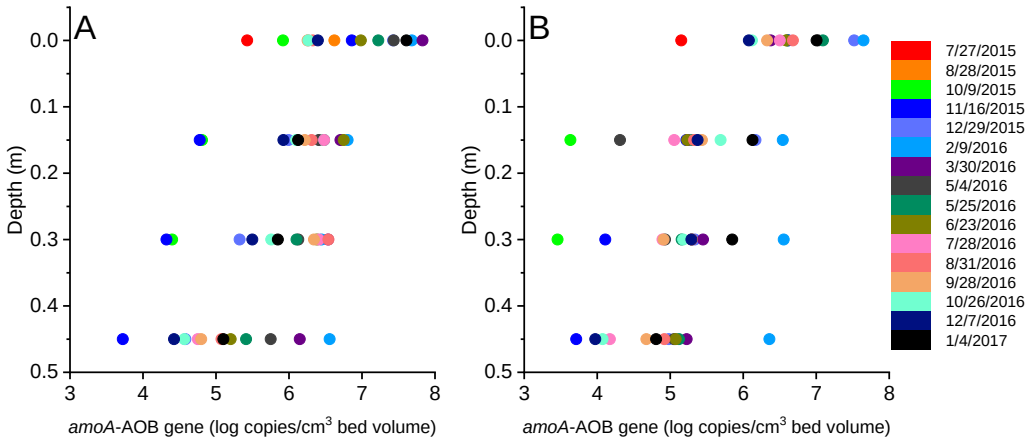

Figure E- 5. *amoA*-AOB gene concentration on media samples collected from four different depth (tap layer, 0.15 m, 0.30 m, and 0.45 m) in the pilot-scale GAC-sand (**A**) and anthracite-sand (**B**) biofilters. Different color represents the sampling date.

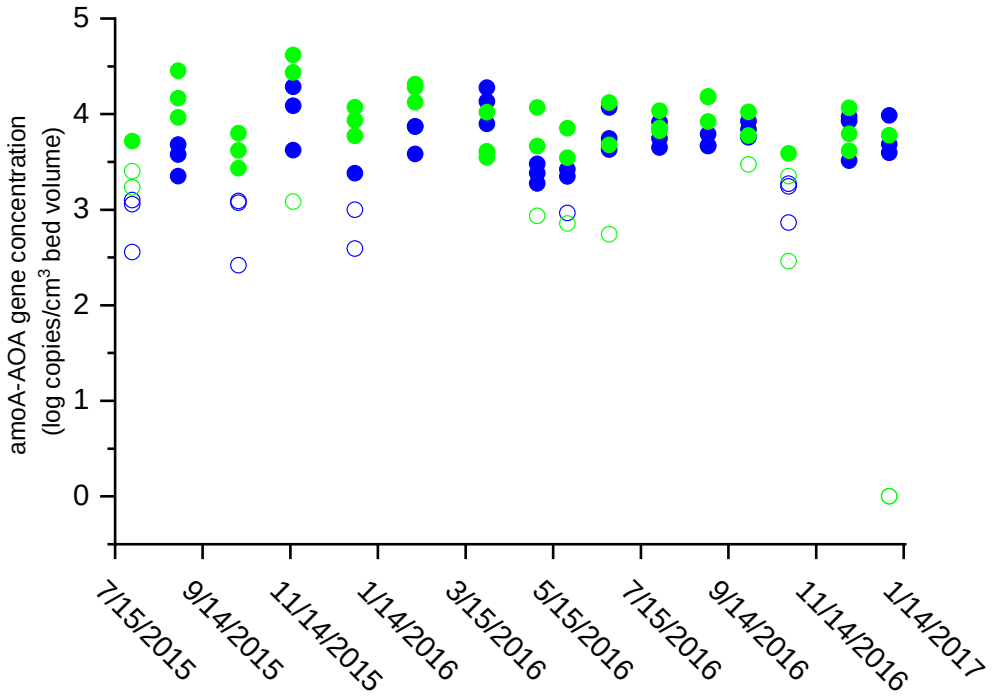

Figure E- 6. *amoA*-AOA gene concentration on media samples collected the top layer of the pilot-scale GAC-sand (blue) and anthracite-sand (green) biofilters. Open symbol represents that the concentration is below the CL.

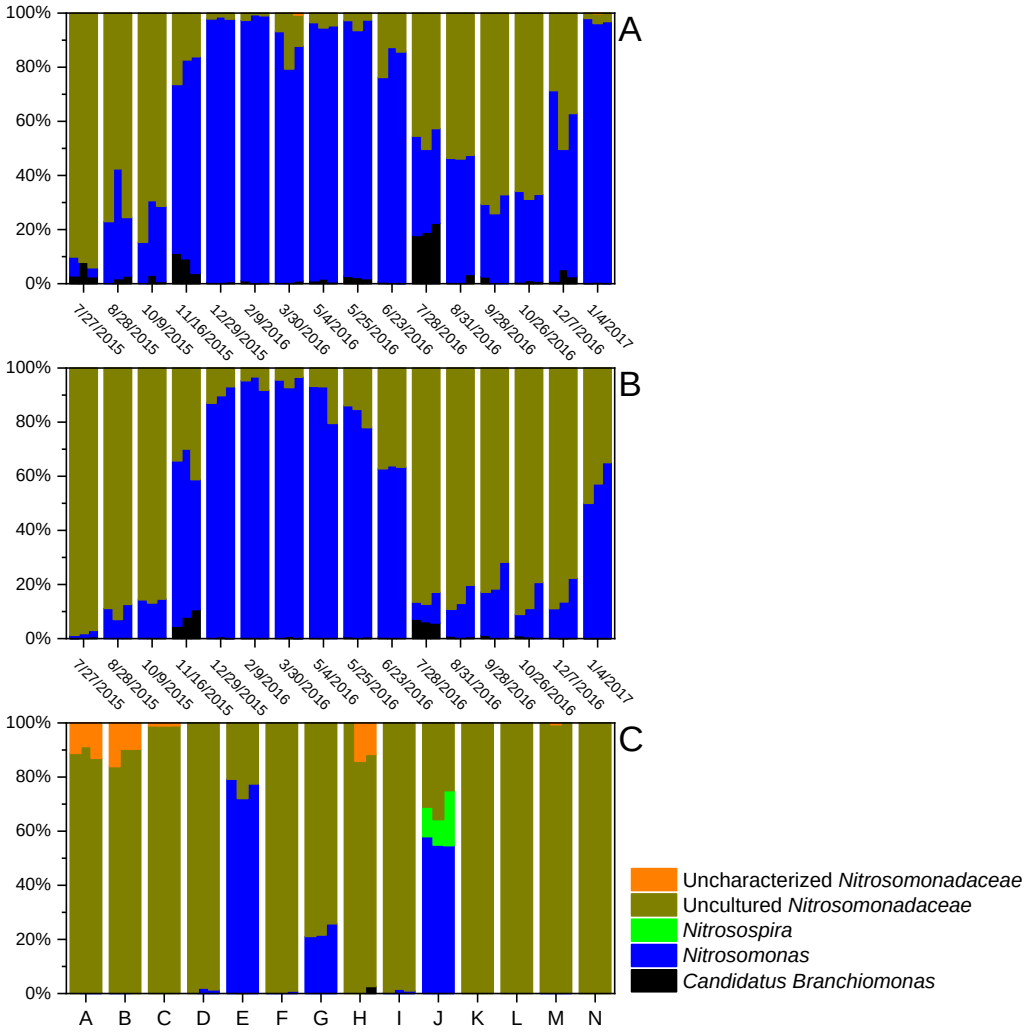

Figure E- 7. Relative contributions of known AOB-like ASVs in the pilot-scale GAC-sand biofilters (**A**), pilot-scale anthracite-sand biofilter (**B**), and full-scale biofilter (**C**) media samples.

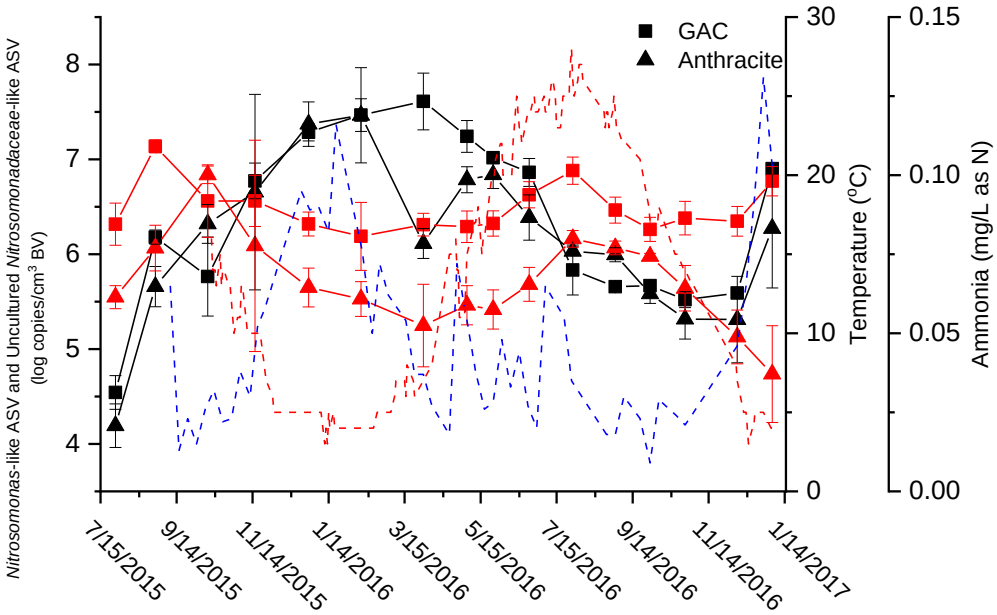

Figure E- 8. Concentrations of *Nitrosomonas*-like (black) and uncultured *Nitrosomonadaceae*-like (red) ASV in GAC and anthracite media samples collected from the pilot-scale biofilters. The red dashed line represents the water temperature. The blue dashed line represents the total ammonia concentration in the filter influent. The error bars represent one standard deviation from mean of triplicate values.

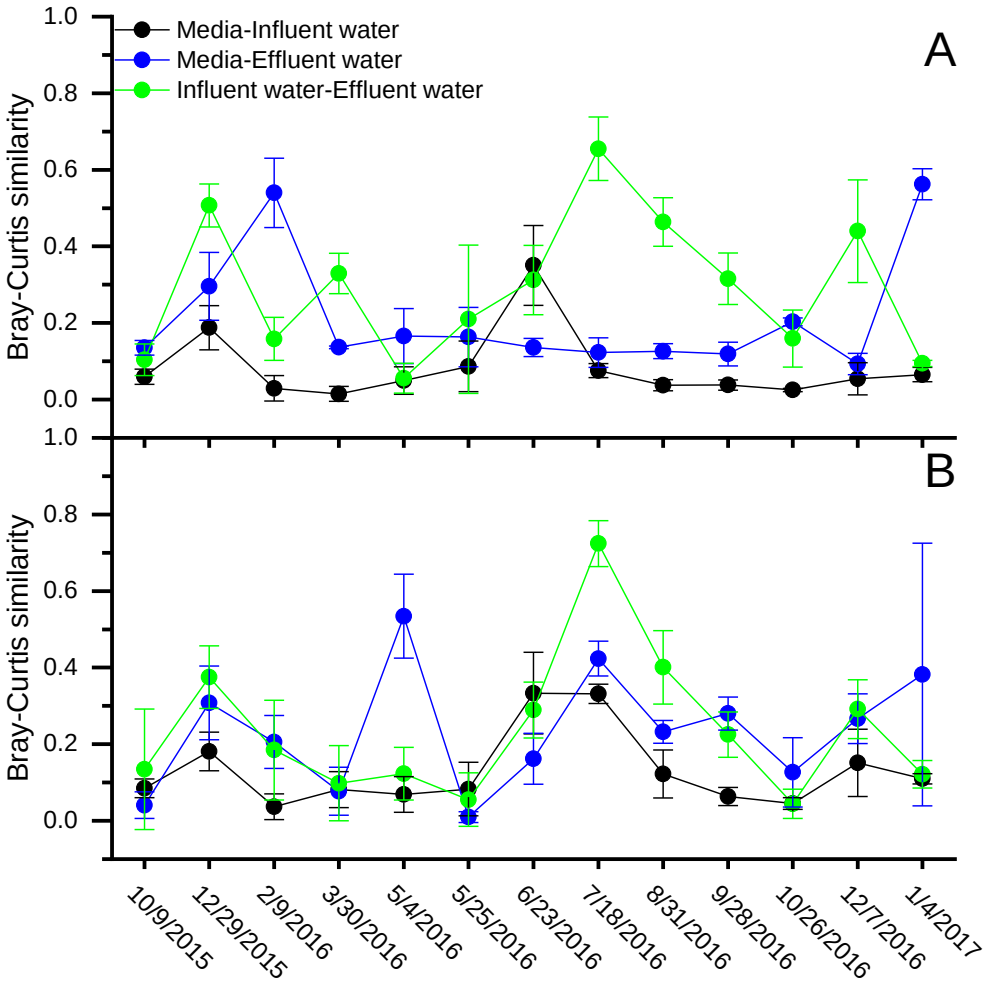

Figure E- 9. Bray-Curtis similarity in ammonia oxidizing bacterial communities between filter influent, filter media and filter effluent collected from the pilot-scale GAC-sand (**A**) and anthracite-sand (**B**) biofilters.

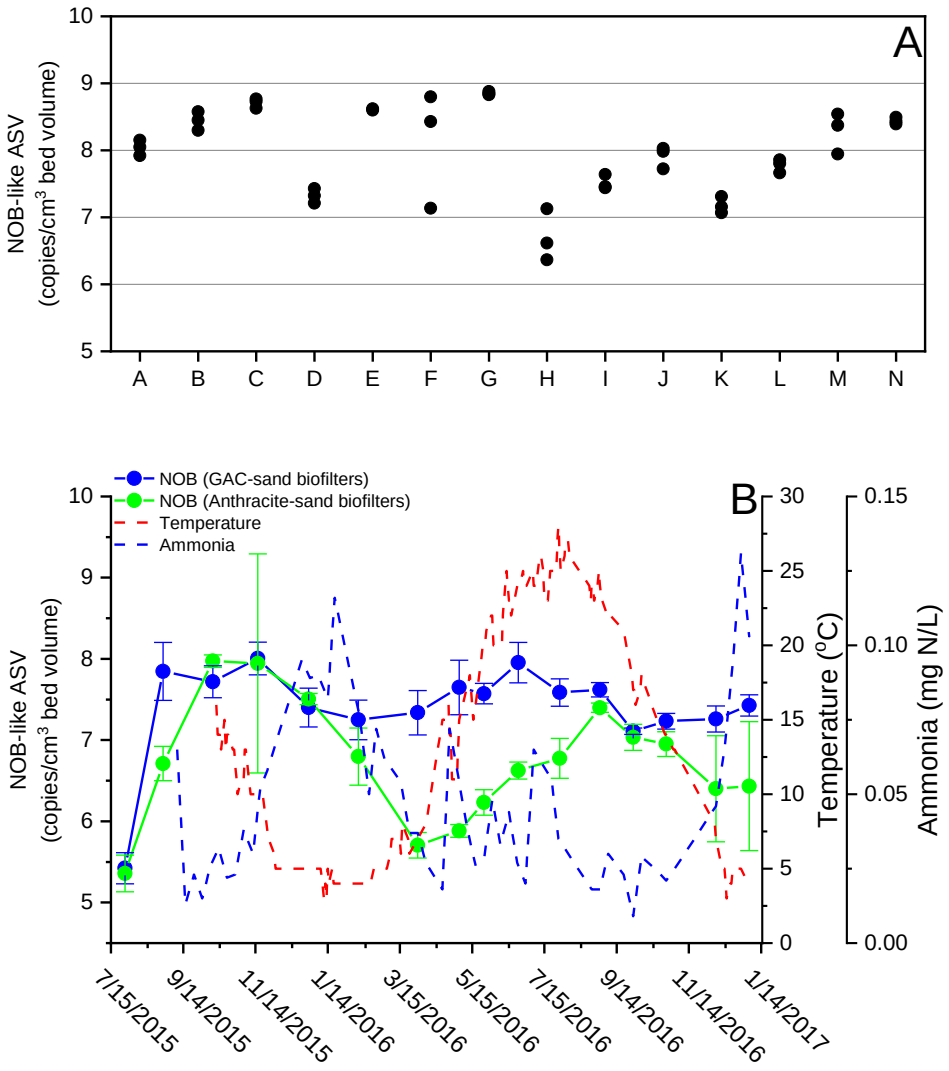

Figure E- 10. Concentrations of NOB-like ASV in the full-scale biofilters (**A**) the pilot-scale biofilters (**B**). X-axis in **A** is the biofilter ID. The error bars in **B** represent the standard deviation computed from triplicate samples.

Table E- 1. Information of primer sequences and gBlock standards for *amoA*-AOB gene and *amoA*-AOA gene in the qPCR assay.

Gene	Primer sequence (5' → 3')	gBlock
amoA-AOB	**F:** TCAGTAGCYGACTACACMGG **R:** CTTTAACATAGTAGAAAGCGG	ATACATGACACCAATAGAGCCTCAAG TAGGAAAGTTCTATAACAGTCCAGTA GCACTGGGTGCAGGTGCTGGTGCAG TATTGTCAGTTACGTTTACAGCATTAG GTTGTAAGCTGAACACTTGGACATACA GCTGGATGGCCGCTTGGTCAGAAGGC GTACTGTTATCAGTAGCTGACTACACA GGTTTCCTGTACGTACGTACCGGTACG CCTGAATATGTTCGCCTGATTGAGCAA GGCTCACTGCGTACCTTTGGTGGACAC ACCACAGTGATTGCAGCGTTCTTCGCA GCGTTCGTATCCATGTTGATGTTCTGC GTATGGTGGTACTTTGGCAAACTGTAC TGCACCGCTTTCTACTATGTTAAAGG CGAAAGAGGACGC
amoA-AOB	**F:** ATA GAG CCT CAA GTA GGA AAG TTC TA **R:** CCA AGC GGC CAT CCA GCT GTA TGT CC	